Oxford
LINGUISTICS

The Phonology of Catalan

Max W. Wheeler

THE PHONOLOGY OF THE WORLD'S LANGUAGES

The Phonology of Catalan

THE PHONOLOGY OF THE WORLD'S LANGUAGES

General Editor: Jacques Durand

Published

The Phonology of Danish
Hans Basbøll

The Phonology of Dutch
Geert Booij

The Phonology of Standard Chinese
San Duanmu

The Phonology of English
Michael Hammond

The Phonology of Norwegian
Gjert Kristoffersen

The Phonology of Portuguese
Maria Helena Mateus and Ernesto d'Andrade

The Phonology and Morphology of Kimatuumbi
David Odden

The Lexical Phonology of Slovak
Jerzy Rubach

The Phonology of Hungarian
Péter Siptár and Miklós Törkenczy

The Phonology of Mongolian
Jan-Olof Svantesson, Anna Tsendina, Anastasia Karlsson, and Vivan Franzén

The Phonology of Armenian
Bert Vaux

The Phonology and Morphology of Arabic
Janet Watson

The Phonology of Catalan
Max W. Wheeler

The Phonology of German
Richard Wiese

In preparation

The Phonology of Tamil
Prathima Christdas

The Phonology of Polish
Edmund Gussman

The Phonology of Spanish
Iggy Roca

THE PHONOLOGY OF CATALAN

Max W. Wheeler

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Cape Town Dar es Salaam Hong Kong Karachi Kuala Lumpur
Madrid Melbourne Mexico City Nairobi New Delhi Shanghai Taipei Toronto

With offices in

Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal
Singapore South Korea Switzerland Thailand Turkey Ukraine Vietnam

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Max W. Wheeler 2005

The moral right of the author has been asserted

Database right Oxford University Press (maker)

First published 2005

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

British Library Cataloguing in Publication Data
Data available

Library of Congress Cataloging in Publication Data

Wheeler, Max
The phonology of Catalan / Max W. Wheeler
p.cm.
Includes bibliographical references
ISBN 0-19-925814-7 (alk.paper)
1. Catalan language–Phonology. I. Title

PC3828.W44 2005
449′.915–dc22 2004066291

Typeset by SPI Publisher Services, Pondicherry, India

Printed in Great Britain
on acid-free paper by
Biddles Ltd., King's Lynn

ISBN 0-19-925814-7 978-0-19-925814-7

1 3 5 7 9 10 8 6 4 2

CONTENTS

Preface x

1 INTRODUCTION 1
1.1 The Catalan language: its varieties and standard(s) 1
1.2 Theoretical assumptions and framework 3
1.3 Outline of contents 5
1.4 Guidance notes 6
1.4.1 Orthography 6
1.4.2 Transcriptions and glosses 8
1.4.3 Abbreviations in glosses 8
1.4.4 A note on Catalan surnames 8

2 PHONOLOGICAL PRIMITIVES AND SEGMENT INVENTORIES 10
2.1 Consonant inventory and contrasts 10
2.1.1 Plosives 10
2.1.2 Affricates 11
2.1.3 Fricatives 13
2.1.3.1 *Labiodental fricative* 13
2.1.3.2 *Alveolo-palatal fricatives* 13
2.1.3.3 *Relation between alveolo-palatal fricatives and affricates* 15
2.1.3.4 *Interpretations of* [ʃ] 22
2.1.3.5 *Dorso-palatal fricative* [j̠] 22
2.1.3.6 *Varieties without voiced strident phonemes* 23
2.1.4 Nasals 24
2.1.5 Glides (semivowels) 24
2.1.6 Rhotics 24
2.1.6.1 *Distribution of rhotic types* 24
2.1.6.2 *Interpretations of rhotic distribution* 26
2.1.7 Laterals 34
2.1.7.1 *Alveolar lateral* 34
2.1.7.2 *Alveolo-palatal lateral* 34
2.1.8 Place in coronals 36
2.1.9 Contrasting geminate consonants 36

2.2 Vowel inventory 37
2.2.1 Stressed vowel contrasts 37
2.2.1.1 *Gaps in the distribution of mid front vowels in central Catalan* 38
2.2.1.2 *Mascaró's law* 41
2.2.1.3 *Instability in the lexical incidence of half-close and half-open mid vowels* 45
2.3 Vowel reduction 52
2.3.1 Introduction 52
2.3.2 Vowel reduction in western Catalan 52
2.3.3 Vowel reduction in Majorcan 53
2.3.4 Vowel reduction in eastern Catalan, except Majorcan 54
2.3.5 The interpretation of Catalan vowel reduction 55
2.3.6 Excluding [ɛ, ɔ] from unstressed syllables 56
2.3.7 Basic vowel reduction in eastern Catalan 57
2.3.8 Basic vowel reduction in Majorcan 60
2.3.9 Exceptions to eastern Catalan vowel reduction, 1: *[əˈa], *[ˈaə], *[əə] 61
2.3.10 Exceptions to eastern Catalan vowel reduction, 2: adjacent mid vowels 65
2.3.11 Exceptions to eastern Catalan vowel reduction, 3: post-tonic [e], [o] 70
2.3.12 Further non-reduction of unstressed [e] in Majorca 73

3 SYLLABLE STRUCTURE 78
3.1 Onsets 78
3.1.1 Onset-driven resyllabification 84
3.2 The syllabification of vocoids 88
3.2.1 Introduction 88
3.2.2 Default syllabification of high vocoids within words 88
3.2.3 High vocoids following other vocoids 90
3.2.3.1 *Stressed vowel + high vowel hiatus* 91
3.2.3.2 *Hiatus before 'underlying' stress* 92
3.2.3.3 *Unstressed vowel followed by unstressed high vocoid* 96
3.2.4 High vocoids preceding other vocoids 99
3.2.4.1 *Labio-velar obstruents?* 100
3.2.4.2 *Other post-consonantal high vocoids followed by non-high vocoids* 102
3.2.4.3 *Vocoid sequences of rising sonority in relation to word stress* 103
3.2.4.4 *High vocoid followed by stressed non-high vocoid in initial syllable* 105
3.2.4.5 *Word-initial pretonic sequences of rising sonority* 109

3.2.4.6 *Initial pre-pretonic sequences of rising sonority* 109
3.2.4.7 *Non-initial sequences of rising sonority, with second element stressed* 110
3.2.4.8 *Pretonic non-initial sequences of rising sonority* 114
3.2.5 High vocoid sequences 120
3.2.5.1 *Adjacent dissimilar high vocoids* 120
3.2.5.2 *Adjacent similar high vocoids* 121

4 PHRASAL PHONOLOGY: VOWEL SANDHI 124
4.1 Introduction 124
4.2 Stressed vowel followed by unstressed vowel 125
4.2.1 Stressed vowel followed by an unstressed high vowel 125
4.2.2 Stressed vowel followed by an unstressed non-high vowel 129
4.3 Unstressed vowel followed by stressed vowel 132
4.4 Contact between unstressed vowels 135
4.5 Observations on multi-vocoid sequences 140
4.6 Summary of evidence for constraint rankings for vowel sandhi 144

5 CODA VOICING NEUTRALIZATION AND ASSIMILATION 145
5.1 Distribution of voiced and voiceless obstruents 145
5.2 'Incomplete' neutralization 149
5.3 Accounting for voicing neutralization and assimilation 151
5.4 Towards an OT account of voicing neutralization in Catalan 157
5.5 Word-final sibilants 162
5.6 Appendix: Steriade's 'P-map' approach to voicing neutralization 164

6 CODA PLACE AND MANNER ASSIMILATION AND NEUTRALIZATION 166
6.1 Introduction 166
6.2 Morpheme-internal coda–C clusters 168
6.2.1 Heterorganic nasal coda clusters 173
6.2.2 Heterorganic denti-alveolar–C clusters 173
6.2.3 Labial–alveolo-palatal cluster 174
6.2.4 Alveolo-palatal–C clusters 176
6.2.5 Velar–C clusters 177
6.2.6 Constraints for formal style 177
6.3 Nasal and lateral assimilation and associated assimilations in less formal style 179
6.4 Inter-word coda–C clusters 180
6.4.1 Inter-word coda–C clusters in less formal styles 183

6.4.2 Summary of style-reflecting place assimilations in inter-word consonantal contacts 185
6.4.3 Heterorganic major place inter-word contacts 185
6.5 Minor place assimilation 186
6.5.1 Survey of cluster realizations 190
6.5.2 Denti-alveolar contacts 192
6.5.3 Alveolo-palatal onsets 197
6.5.4 High DAC contacts 201
6.5.5 Conclusions and Summary 204
6.6 Consonantal contact in Majorcan: place and manner 207
6.6.1 Labial coda inputs 211
6.6.2 Alveolar coda inputs 214
6.6.3 Alveolo-palatal coda inputs 215

7 CLUSTER REDUCTION 220
7.1 Introduction 220
7.2 Cluster reduction in stems 221
7.3 Cluster reduction in stems with a consonantal suffix 227
7.4 Cluster reduction in uninflected stems followed by vowel-initial words 235
7.5 Cluster reduction in uninflected stems followed by consonant-initial words 237
7.6 Cluster reduction in inflected stems followed by vowel-initial words 240
7.7 Cluster reduction in inflected stems followed by consonant-initial words 244
7.8 Summary of constraint rankings relating to word-final clusters 249

8 EPENTHESIS AND OTHER SONORITY-RELATED PHENOMENA 250
8.1 Introduction 250
8.2 Non-edge epenthesis 252
8.3 Codas and epenthesis 257
8.4 Sibilants and epenthesis 263
8.5 Minimum sonority distance and voiced stop gemination 265
8.6 Balearic verb forms with coda clusters violating the sonority sequencing principle 269

9 STRESS, FEET, AND PHRASES 276
9.1 Word stress: introduction 276
9.2 Unstressed words 277

9.3 Compounds and complex words 279
9.3.1 Compounds 279
9.3.2 Complex words with affixes 282
9.4 Verb stress 284
9.5 Prosodic word constraints 288
9.6 Rhythmic constraints above the word: a provisional account 297
Appendix: evaluation of alternative foot and colon parses for word types in (29–31) 306

10 WORD PHONOLOGY AND PHONOLOGICALLY CONDITIONED ALLOMORPHY 310
10.1 Coda/Onset alternations: lenition 310
10.1.1 Introduction 310
10.1.2 Fricatives or approximants 312
10.1.3 Lenition and the theory of effort 314
10.1.4 Lenition and the enhancement of contrast 316
10.1.5 Variation in Catalan lenition 317
10.1.6 Further interpretation of the hierarchy of leniting contexts in Catalan 324
10.1.7 Remaining issues 326
10.2 Consonant/zero alternations 327
10.2.1 N/zero alternation 327
10.2.2 R/zero alternation 333
10.3 Alternations involving stem-final labials 338

11 THE SYLLABIFICATION OF PRONOMINAL CLITICS 341
11.1 Introduction 341
11.2 Single clitics 342
11.3 Clitic clusters 345
11.4 Syllabification constraints affecting clitics and clitic clusters 348
11.4.1 Epenthesis 348
11.4.2 Deletion of consonants and other issues in the syllabification of pronominal clitics 355
11.4.3 Non-edge epenthesis in verb–clitic sequences 363
11.5 Epenthesis in clitic clusters to avoid geminate sibilants 366
11.6 Summary of constraint ranking 369

References 371

Index 382

PREFACE

I first worked on Catalan phonology between 1969 and 1974, work that resulted in an Oxford doctoral thesis, later published in book form (Wheeler 1975/9). That piece of research applied to standard Barcelona Catalan pronunciation the 'classical' Generative Phonology approach of Chomsky & Halle's *The Sound Pattern of English* (1968). What has changed since then to justify a new book on Catalan phonology from the same author? Not fundamentally the language itself. The present book does, however, look at a much wider range of varieties of Catalan than the earlier one, though without aspiring to comprehensiveness on dialect phonology. I also deal explicitly with style and other social variation throughout, not least because this aspect has received little attention in works on Catalan published in English. Since 1980, in particular, a considerable amount of research on Catalan phonology has appeared, much of it in Catalan, and not all of it in publications readily accessible outside the Catalan-speaking area. So one of my aims is to bring the insights of some of this work to the attention of English-using phonologists. There is no doubt, though, that the main excuse for my devoting a new book to the subject is the revolution that has taken place within phonological theory since Smolensky & Prince's work on Optimality became available in 1993. That is the major reason why a phonology of Catalan looks, and should look, very different at the start of the twenty-first century from what such a thing looked like a generation previously. Simply put, we now ask very different questions about the same (or comparable) facts. And we look at the phonology of a language not 'on its own terms' but in the light of as much as we can know and understand of the phonology of other languages.

The writing of this book has been facilitated by the generous assistance of phonologists in the Catalan-speaking countries. I am especially grateful first to those who have supplied me with copies of their books, offprints, pre-prints and typescripts, or answered questions: Eulàlia Bonet, Teresa Cabré, Jordi Colomina, Nicolau Dols, Jesús Jiménez, Sílvia Llach, Maria-Rosa Lloret, Joan Mascaró, Blanca Palmada, Josep Pi-Mallarach, Clàudia Pons, Pilar Prieto, Daniel Recasens, and Pep Serra; and secondly to those who have been kind enough to offer comments, criticisms, and suggestions on drafts of parts of the text, including the participants in the Barcelona Phonology Seminar in May 2003, and in particular Teresa Cabré, Nicolau Dols, Pere Grimalt, Maria-Rosa Lloret, Clàudia Pons, Pilar Prieto, and Daniel Recasens. I am alone responsible for the use I have made of their observations. I would also like to express my thanks to Iggy Roca, who read the whole of the text on behalf of Oxford University Press and suggested

clarifications both of thought and expression. The book is the better for his input, even though I declined some of his recommendations.

Work on this project was supported by a period of research leave (2001–2) from the University of Sussex, partly funded by a grant from the Arts and Humanities Research Board.

Max Wheeler
Oldlands, May 2005

This page intentionally left blank

1

INTRODUCTION

1.1 THE CATALAN LANGUAGE: ITS VARIETIES AND STANDARD(S)

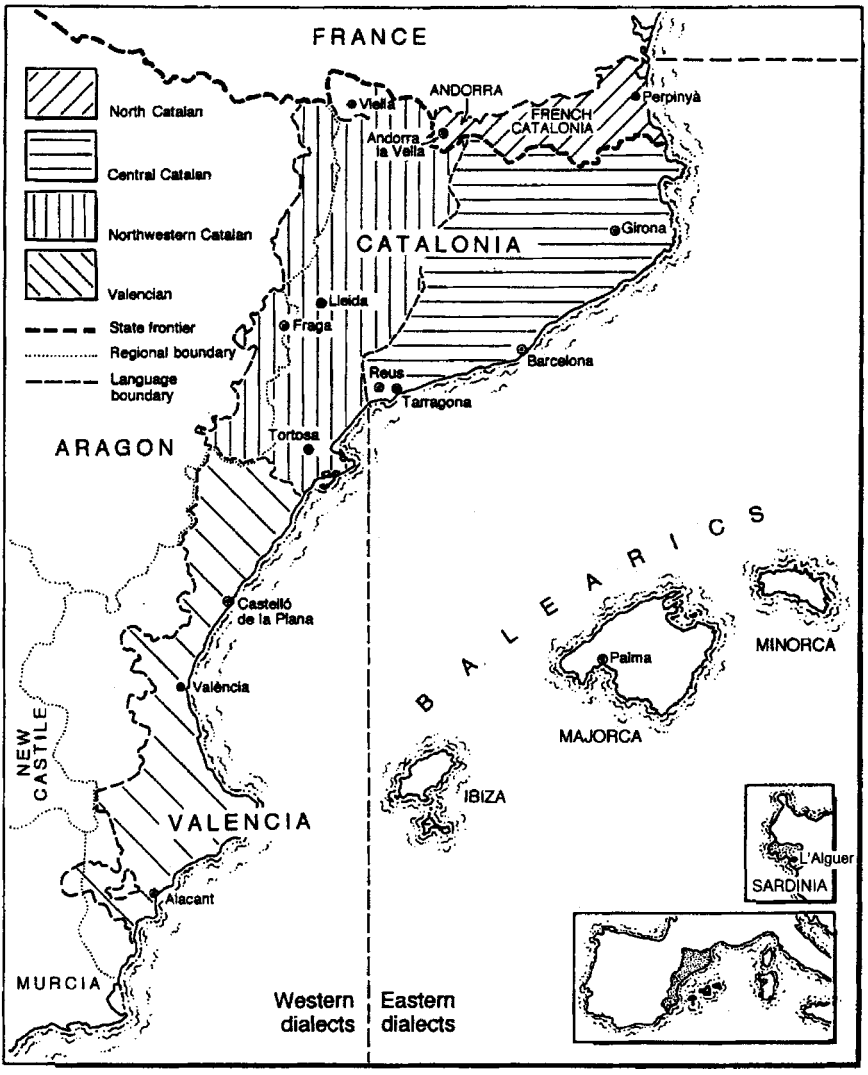
Catalan is a member of the Romance family of languages—modern forms of the Latin language which spread from Rome to much of Europe in the wake of the Roman empire around the beginning of our era.

The territories where Catalan is natively spoken cover almost 60,000 km^2, of which 96.2 per cent lies within Spain. They are:

(a) the Principality of Andorra
(b) in France: almost all of the *département* of Pyrénées-Orientales
(c) in Spain:
Catalonia (under the autonomous government, the *Generalitat*, of Catalonia), except for the Gascon-speaking Val d'Aran
the eastern fringe of Aragon
most of the *Comunitat Valenciana*, excepting some regions in the west and south which have been Aragonese/Spanish-speaking since at least the eighteenth century
El Carxe, a small area of the province of Murcia, settled in the nineteenth century
the Balearic Islands: Majorca and Minorca (Balearic *stricto sensu*), Ibiza (Catalan *Eivissa*) and Formentera (strictly *Illes Pitiüses*)
(d) in Italy: the port of Alghero (Catalan *L'Alguer*) in Sardinia

As its geographical position might suggest, Catalan shares several features with its nearest Romance neighbours—Italian, Sardinian, Occitan, and Spanish—while being distinct in important respects from all of them. For the most part the boundaries between Catalan and neighbouring languages are quite sharp, but the dialect of the valley of Capcir is transitional between Catalan and Occitan, and the dialects of Alta Ribagorça in the Aragonese Pyrenees are transitional between Catalan and Aragonese. The map on page 2 shows where Catalan is spoken and identifies the main dialect divisions.

Though there are significant dialect differences in Catalan, the dialects are to a high degree mutually intelligible. They are conventionally divided into two groups, on the basis of differences in phonology as well as in some features of



MAP 1.1. Catalan-speaking areas and dialects (from Wheeler et al. 1999: xviii)

verb morphology; there are some interesting lexical differences, too. The eastern dialect group includes North Catalan or *rossellonès* (in French Catalonia), central Catalan (in the eastern part of Catalonia), Balearic, and *alguerès* (in Alghero/*Alguer*). The western group consists of north-western Catalan (western and southern Catalonia and eastern Aragon) and Valencian. The main 'diagnostic heterogloss' distinguishing the two major dialect groups involves vowel reduction in unstressed syllables (§2.3): in the eastern dialects /a/ is pronounced [ə] in

unstressed syllables, and, with some exceptions, /e/ and /ɛ/ are also reduced to [ə], while /o/ and /ɔ/ are reduced to [u].

A conservative estimate of the number of native speakers of Catalan is about 6.5 million (based on the 1991 census in Catalonia, Valencia, and the Balearics, with estimates for the remaining territories). Within Spain, a further 3 million claim to understand Catalan. Partly as a result of the incorporation of Catalan locally into the education system, there are within Spain increasing numbers of second-language speakers. Virtually all speakers of Catalan are bilingual, also using the major language of the state they live in. (Andorrans are bilingual in Spanish or French, or are trilingual.)

1.2 THEORETICAL ASSUMPTIONS AND FRAMEWORK

The present account of Catalan phonology falls squarely within the bounds of current Optimality Theory, whose principles, after a decade or so of research, can now be found in basic texts such as Kager (1999), McCarthy (2002), and McCarthy (2004). As the focus of this book is the facts of Catalan pronunciation, I have avoided taking up opportunities to argue for substantial theoretical innovations. Thus the variety of Optimality Theory reflected here may be called 'conservative' or standard, though many of the specific constraints invoked derive explicitly from recent work by others in the field. By 'conservative' or 'standard' OT, I mean a version with strict ranking of constraints, and with only one level of grammar, enriched by different correspondence constraints—for example, between morphological bases and their derivatives, or between prosodic words and their realization in phonological phrases.

Though in this book I adopt what is intended as an orthodox OT framework, there are certain respects in which I believe standard OT is inadequate as a realistic theory of human speech behaviour. The first of these concerns the problem of modelling variation. The strict ranking requirement entails that all variation reflects different ranking of constraints, which is to say: different grammars. In this respect OT has nothing to say about why variation exists or how it is acquired or diffused. I take it that strict ranking of constraints reflects, in fact, no more than a convenient idealization of the facts, such as will allow an explicit account of categorical phenomena. Such an idealization is a methodological move that resembles the neo-grammarian theory of regular sound change, or the Chomskyan ideally competent speaker-hearer in a homogeneous speech community. Categorical behaviour is certainly part of human language, and it is a reasonable simplification to attempt to grasp the categorical before we explore the variable. But it seems to me likely that the categorical aspect of phonological behaviour results not from constraints strictly ranked by nature, but from constraint weightings, or strengths, that are different enough to produce categorical or quasi-categorical outcomes. I suspect that in a more realistic model constraints are probabilistic; some variation, at least, results from constraint interactions

whose output is strictly indeterminate. Most phonological research, except for that in the Labovian tradition, has looked into those phenomena that are categorical, or so nearly so that a simplification is not a distortion. This is as true of research in Catalan as it is of research into other languages. Most phonological variation has not yet been seriously investigated. In this book two sources of both inter-speaker and intra-speaker variation are hinted at: different degrees of awareness of morphological structure (Chapter 3), and differences of speech rate (Chapter 4 and §9.6).

A second respect in which Optimality Theory assumes a convenient simplification concerns precisely what aspects of knowledge are being modelled. Standard OT aims to be neutral between perception and production. But phonological output representations that are the basis of constraint evaluation correspond rather more directly to what is perceived than to what is articulated. An output representation is, usually, a perceptual representation of phonetic segment types enriched by prosodic elements such as syllable and foot structure. This perceptual bias is another convenient idealization that can become problematic when there may be a mismatch between (intended) production and (perceived) output—some examples occur here in Chapters 3 and 5. A major piece of work in phonological theory that addresses both the issue of perception versus production and the issue of probabilistic constraint ranking or weighting is Boersma (1998).

In the two respects mentioned in the previous paragraphs I believe current Optimality Theory is not yet sufficiently rich to encompass a fully realistic account of phonology. In other respects, however, I believe it is insufficiently restrictive. One such area is that of alignment constraints. According to the model of Generalized Alignment (McCarthy & Prince 1993), either edge of any grammatical or prosodic category can be required to coincide with either edge of any other grammatical or prosodic category. As McCarthy (2002: 46, n. 14) points out, 'alignment constraints have a somewhat ambiguous status in markedness/faithfulness constraint typology'. McCarthy might have said more explicitly that alignment constraints are neither markedness constraints nor faithfulness constraints. I believe they are predominantly stipulative in character, or to put it another way, they differ little, if at all, from phonological *rules*. Almost certainly some alignments are more marked than others,[1] and it is appropriate that constraint rankings should reflect this. This account of Catalan makes considerable use of alignment constraints, though only of the kind that align a grammatical category (Stem, Lexical Word, Affix, etc.) with a prosodic category (Prosodic Word, Syllable, Foot, etc.), or a prosodic category with another at a different level.[2] For aligning one grammatical category with another, or, in principle, one prosodic category with another at the same level, I prefer to

[1] For example, infixation is more marked than prefixation which is more marked than suffixation.

[2] An exception is the ALIGN(Clitic-Verb) constraint in Ch. 11 that is borrowed from Bonet & Lloret (2001, forthcoming).

use CONTIGUITY constraints, which are explicitly correspondence/faithfulness constraints that penalize interruption of input material (see Landman 2002).

I also adopt from standard OT the conventional view of input representations, where the elements are something like classical phonemes: contrastive elements with inherent phonetic content. In OT, though, some aspects of the input may be indeterminate, or irrelevant. It is the task of the constraints to ensure grammatical outputs. Some recent work suggests a more subtle approach, including the idea that underlying representations can be eliminated, with contrasting elements determined by a ranking of markedness constraints (Burzio 1999, Flemming 1995; Padgett 1997; 2001; Ní Chiosáin & Padgett 2001), though McCarthy (2003a) gives good reasons to believe that underlying representations are not, in fact, entirely determined by contrast principles, though they are no doubt motivated by such principles. On related lines Łubowicz (2003) proposes a new type of PRESERVECONTRAST constraints. I think these ideas promise developments towards greater realism in phonology, and I regret that space limitations have not allowed me to do more than hint at these questions (Chapters 8, 10) in the context of Catalan.

The majority of work published in phonological Optimality Theory has been driven by a research agenda that focuses on theoretical or typological issues. And this is entirely appropriate. With such a perspective, the analyst is free to choose the data which best illustrate some theoretical issue, or the data which serve to fill out a typology. With such objectives it is legitimate to ignore data that complicate, without bearing critically upon, the issues addressed. The constraints on a phonologist attempting to deal with the whole range of phonological phenomena in one language, as in this *Phonology of Catalan*, lead to a somewhat different relation between data and theory. Responsibility to the data in this context prompts the phonologist to an attempt to be comprehensive, and this involves the exposition of a range of data whose interpretation may be more provisional, or controversial, than would be appropriate in an issue-driven paper. The reader of a language-oriented work of phonology is entitled to know not just the facts which can be explained by current theories but also those which may seem problematic, as well as those which may not illuminate topics of current theoretical concern. Of course, many facts of Catalan pronunciation, as is the case with many elements of any language, are not yet 'known' to the community of linguistics, but I have attempted to include mention, and in most cases some interpretation, of what is 'known'. Where my accounts of the phenomena are speculative, they may be regarded as standing in for better accounts to be hoped for in future research.

1.3 OUTLINE OF CONTENTS

This book deals with the distinctive phonological elements of Catalan: consonants (§2.1), vowels (§2.2), and word stress (Chapter 9), and the ways in which these elements are realized in different contexts. Contextually governed

realization includes phenomena such as assimilation (Chapters 5 and 6), lenition (§10.1), gemination (§2.1.2–3, §8.5), and syllabification, both of high vocoids (Chapters 3 and 4) and of pronominal clitics whose underlying form is consonantal (Chapter 11). Further processes involve significant contextual neutralization of contrast, such as vowel reduction (§2.3), assimilation of the features of voicing, place, and manner in consonant clusters (Chapters 5 and 6), and several types of neutralization of contrast with zero: consonant cluster reduction (Chapter 7), vowel elision (Chapter 4), and epenthesis (Chapters 8 and 11). As is to be expected, cluster reduction, elision/fusion, and epenthesis are largely driven by marked syllable structure constraints, though elision and glide formation also respond in part to rhythmic constraints. A considerable proportion of the book deals with 'post-lexical' phenomena, in particular, sandhi, which is a major focus in Chapters 4–8. Relatively little space is given to morphophonemic phenomena where alternations are not governed by markedness constraints, though such are in view in §10.2–3.[3] Beyond the phenomena treated in those sections, 'lexicalized' phonology seems not to be very important in modern Catalan, or at any rate to be more illuminatingly discussed from a specifically morphological perspective; see Wheeler (2002), for example, on irregular stem alternation in verb morphology. Among suprasegmental phenomena, apart from word stress (Chapter 9) I give some attention, in §9.6, to the issue of rhythm, though the account of it I offer is indicative rather than comprehensive. I make no attempt to deal with Catalan intonation, though a substantial amount of descriptive and analytical work has been carried out by Catalan linguists.[4]

The dialect of Catalan about which most has been published is central Catalan, the variety of northern and eastern Catalonia, including the city of Barcelona. This is also the dialect I know best, and it consequently receives the most detailed treatment here. I also provide a considerable amount of information about western varieties, including Valencian, and about Balearic, especially Majorcan, insofar as these dialects differ systematically from central Catalan. However, I do not touch at all on the Catalan dialect of Alghero (Sardinia), whose phonology is markedly, and interestingly, different from the rest.

1.4 GUIDANCE NOTES

1.4.1 *Orthography*

Catalan orthography is systematic and largely phonologically based. The roman alphabet is extended with the following symbols and digraphs:

[3] Phonological realization of inflectional morphology is treated at length in Wheeler (1975/9), in the framework of Generative Phonology current at that time.

[4] Important sources are Astruc (2003), Bonet (1984), Estebas-Vilaplana (2000; 2003), Payà (2003), Pradilla (1998), Prieto (1996; 1997; 1999; 2002a; 2002b), Recasens (1978), Salcioli Guidi (1988), Tió (1987).

ç	(*ce trencada*) pronounced /s/, as in *feliç* [fəˈlis] 'happy'
gu	pronounced /ɡ/ ([g] ∼ [ɣ]) before *i* and *e*, as in *guerra* [ˈɡɛrə] 'war' (*gu* is /ɡw/ before other vowels, as in *guant* [ˈɡwan] 'glove')
ig	pronounced [t͡ʃ] in final position, as in *raig* [rat͡ʃ] 'trickle'
ix	pronounced /ʃ/ ([jʃ] in some dialects), as in *caixa* [ˈkaʃə] 'box'
l·l	(*ela geminada*) normatively /ll/ but usually /l/, as in *novel·la* [nuˈβɛlə] 'novel'
ny	pronounced /ɲ/ as in *Catalunya* [kətəˈluɲə] 'Catalonia'
qu	pronounced /k/ before *i* and *e*, as in *qui* [ˈki] 'who' (*qu* is /kw/ before other vowels, as in *quatre* [ˈkwatrə] 'four')
ss	pronounced /s/, as in *grossa* [ˈɡrɔsə] 'big.F' (intervocalic *s* corresponds to /z/, as in *casa* [ˈkazə] 'house')
tg and *tj*	pronounced [d͡dʒ], as in *fetge* [ˈfed͡dʒə] 'liver', *mitjó* [mid͡ˈdʒo] 'sock'
tx	pronounced [t͡ʃ], as in *despatx* [dəsˈpat͡ʃ] 'office'
tz	pronounced [d͡dz], as in *dotze* [ˈdod͡dzə] 'twelve'

Several letters and digraphs have contextually conditioned pronunciations, leading often to regular spelling alternations in morphologically related forms:

c	before *e* or *i*, pronounced /s/, corresponds to *ç* in other positions, hence *feliç* 'happy.M' – *felices* 'happy.F.PL', *caço* 'hunt.1SG.PRS.IND' – *caces* 'hunt.2SG.PRS.IND'
g	before *e* or *i*, pronounced /ʒ/, corresponds to *j* in other positions, hence *envejar* 'envy.INF' – *envegen* 'envy.3PL.PRS.IND'
final *g*	after stressed *i*, and final *ig* after other stressed vowels, pronounced [t͡ʃ], corresponds to *j* ∼ *g* or to *tj* ∼ *tg* in other positions, hence *boig* [ˈbɔt͡ʃ] 'mad.M' – *boja* 'mad.F' – *boges* 'mad.F.PL'; *desig* [dəˈzit͡ʃ] 'wish' – *desitjar* 'wish.INF' *desitgem* 'wish.1PL.PRS.IND'
gu	before *e* or *i*, pronounced /ɡ/, corresponds to *g* in other positions, hence *botiga* 'shop', plural *botigues*, *llarg* 'long.M' – *llargues* 'long.F.PL'
gü	before *e* or *i*, pronounced /ɡw/, corresponds to *gu* in other positions, hence *llengua* 'language', plural *llengües*
qu	before *e* or *i*, pronounced /k/, corresponds to *c* in other positions, hence *vaca* 'cow' plural *vaques*; *sac* 'bag' – *saquet* 'bag.DIM'
qü	before *e* or *i*, pronounced /kw/, corresponds to *qu* in other positions, hence *obliqua* 'oblique.F' – *obliqües* 'oblique.F.PL'
x	is pronounced [ʃ] ∼ [t͡ʃ] initially and in onsets after a consonant, [ʃ] after *i*, otherwise [gz] before stress, [ks] after, hence *xarxa* [ˈʃarʃə] 'net', *guix* [ˈɡiʃ] 'chalk', *exacte* [əɡˈzaktə] 'exact', *fax* [ˈfaks] 'fax'

Word-final devoicing of non-strident obstruents is reflected in the orthography immediately after a stressed vowel, but not generally elsewhere, hence:

dic	[ˈdik] 'say.1SG.PRS.IND' – *diguem* [diˈɣɛm] 'say.1PL.PRS.SUBJ', *salut* [səˈlut] 'greeting' – *saludar* [səluˈða] 'greet.INF', *llarg* [ˈʎark] 'long.M' –

llarga [ˈʎaɾɣə] 'long.F', *líquid* [ˈlikit] 'liquid.M' – *líquida* [ˈlikiðə] 'liquid.F'.

The diaeresis ¨ has two functions: in *qü* and *gü* to show [w] pronunciation (see above), and in *ï* and *ü* following a vowel, to indicate hiatus rather than a glide, as in *agraïa* [əɣɾə.ˈi.ə] 'thank.1/3SG.PST.IMPF' (contrast *atzagaia* [ədz͡əˈɣajə] 'javelin'), *peüc* [pəˈuk] 'bed sock' (contrast *meuca* [ˈmɛwkə] 'prostitute').

Acute and grave accents, where used, indicate the position of lexical stress; over *e* and *o* acute accents represent half-closed vowels [e] and [o], grave accents represent half-open [ɛ] and [ɔ].

1.4.2 *Transcriptions and glosses*

Transcriptions unless otherwise stated represent central Catalan pronunciation, and are generally broad phonetic, though more phonetic detail, or syllable structure, is included in certain chapters where greater detail is in focus. The symbol ˈ indicates the position of word stress, not necessarily primary phrase stress. Glosses, as usual, are primarily intended to identify the lexical item or grammatical form in question.

1.4.3 *Abbreviations in glosses*

For abbreviations in the glosses I attempt to follow the standard model used in the European Typology project, and set out in Croft (2003). The following list contains those used in this book.

≺	less harmonic than	DO	direct object	PEJ	pejorative
≻	more harmonic than	F	feminine	PFCT	perfect
1	first person	FUT	future	PL	plural
2	second person	IMP	imperative	POL	polite
3	third person	IMPF	imperfective	PRF	perfective
ACC	accusative	IND	indicative	PRS	present
AUG	augmentative	INF	infinitive	PRTT	partitive
AUX	auxiliary	INTR	intransitive	PST	past
COMP	complementizer	IO	indirect object	REFL	reflexive
COND	conditional	LOC	locative	RPRN	relative pronoun
DAT	dative	M	masculine	SG	singular
DIM	diminutive	N	neuter	SUBJ	subjunctive

1.4.4 *A note on Catalan surnames*

Within Spain, people officially have two surnames, of which the first is one's father's first surname and the second is one's mother's first surname. Some Catalan writers regularly publish using both surnames, often linked with *i*

'and'. One linguist who does so is Antoni M. Badia i Margarit. Others only ever use their first surname, as does Joan Mascaró. Others again sometimes publish under their first surname alone and sometimes under both, as does Daniel Recasens i Vives. In a case like the last, in the list of References I supply the second surname in square brackets where it is missing from the published text, if only to confirm that the same author is involved. Whether or not an author uses one or both surnames in publications, it is appropriate to identify them in a citation just by the first surname. One may, nonetheless, mention both surnames to distinguish individuals with the same first surname. Thus Montserrat Badia i Cardús is cited as 'Badia i Cardús' to distinguish her from her father, Antoni M. Badia i Margarit. In all cases, alphabetization is by first surname.

2

PHONOLOGICAL PRIMITIVES AND SEGMENT INVENTORIES

2.1 CONSONANT INVENTORY AND CONTRASTS

In this section I introduce the major consonant types found in Catalan (Julià 2002). Table 2.1 sets out the consonant types, according to articulatory criteria. Not all the distinctions in the table are phonemic. The labels for place of articulation correspond to the realization of consonants in the contexts least affected by coarticulation, that is, basically, prevocalic position. A more precise classification of the place of articulation of coronals identifies [t], [d]/[ð] as having both dental and alveolar contact; [n], [l], and [ɾ] are front alveolar; [s]/[z] and [r] are back alveolar; [ʎ] and [ɲ] are front alveolo-palatal;[1] and [ʃ]/[ʒ] and [t͡ʃ]/[d͡ʒ] are back alveolo-palatal (Recasens and Pallarès 2001b). The discussion then focuses on issues of phonemic interpretation, phonemic contrast, and lexical incidence, drawing attention to some interesting differences between Catalan dialects.

2.1.1 *Plosives*

The voiced plosives of all places of articulation are in complementary distribution with the voiced fricatives of the same place: [b] ∼ [β], [d] ∼ [ð], [g] ∼ [ɣ]. Broadly, the fricative allophones are found only in syllable onsets, after continuants; the plosives are found elsewhere; an account of this distribution is offered in §10.1.2. The dorso-palatal stops [c], [ɟ] are found only in some Majorcan varieties, corresponding to the velars of the remaining varieties; see §2.1.3.5. These Majorcan dorso-palatals have velar allophones only before liquids and rounded (back) vowels, but are palatal before other vowels, including low central [a], and in word-final codas, such as *banc* [ˈbanʲc] 'bench' (Dols 2000).

[1] The term is that of Recasens & Pallarès; their interpretation differs from the terminology of the IPA, where [ʃ] and [ʒ] appear under the heading 'post-alveolar', and [ʎ] and [ɲ] under the heading 'palatal', with the implication that 'palatals' are pronounced with contact further back than 'post-alveolars'.

TABLE 2.1 Inventory of major consonant types in Catalan

	Bilabial	Labio-dental	Lamino-dental	Apico-alveolar	Lamino-alveolo-palatal	Dorso-palatal	Velar/Labio-velar
Plosive	p b		t d			c ɟ	k g
Affricate				t͡s d͡z	t͡ʃ d͡ʒ		
Fricative[2]	β	f v	ð	s z	ʃ ʒ	ʝ	ɣ
Nasal	m			n	ɲ		ŋ
Approximant						j	w
Tap				ɾ			
Trill				r			
Lateral				l	ʎ		

2.1.2 *Affricates*

The status of the affricates of Catalan as unit phonemes has been much debated (just in the fifteen years up to 2002, see Mariner 1987; Recasens 1991b: 214; Borràs 1992; Badia i Cardús 1993; 2000; Recasens 1993: 160–61; Jiménez 1996a). I consider in this section only the affricates that contrast with fricatives. Those affricates that are in variation with fricatives are taken up in §2.1.3. Some contrasting examples of affricates are illustrated in (1).

(1) Affricate — Fricative

atxa [ˈat͡ʃə] ~ [ˈat͡ʃt͡ʃə] 'torch' — *aixa* [ˈaʃə] 'adze'

despatx [dəsˈpat͡ʃ] 'office' — *calaix* [kəˈlaʃ] 'drawer'

mitja [ˈmid͡ʒə] ~ [ˈmid͡ʒd͡ʒə] 'stocking' — *pruïja* [pɾuˈiʒə] 'itch' (Val. [pɾuˈid͡ʒa])

In the case of forms like those in (1), the issue in debate has been whether [t͡ʃ] ~ [t͡ʃt͡ʃ] is to be regarded as consisting of (or containing, in the case of geminate pronunciation) a unit affricate /t͡ʃ/, or as a sequence of a stop /t/ followed by a fricative /ʃ/, and likewise for [d͡ʒ] ~ [d͡ʒd͡ʒ]. In Valencia and southern Catalonia there is no voiced fricative [ʒ], but only [d͡ʒ], and affricates occur word-initially where other varieties have fricatives (e.g. *xinxa* 'bedbug', Valencian and southern Catalonia [ˈt͡ʃinʲt͡ʃa], elsewhere [ˈʃinʲʃə]). In these varieties the unit affricate interpretation seems the most likely for the alveolo-palatal affricates in (1), at least (/t͡ʃ/, /d͡ʒ/), inasmuch as they occur in word-initial position where obstruent clusters are not found, as well as medially and finally; see §2.1.3.2. But in the other varieties, which lack word-initial invariant alveolo-palatal affricates, the balance tips towards the biphonemic interpretation of the affricates that contrast with fricatives. The sequence interpretation is also more plausible for the alveolar affricates [t͡s] and [d͡z] ~ [d͡zd͡z] which are absent word-initially in all varieties. For example, intervocalic affricates are predominantly long, particularly the voiced

[2] See §§2.1.3.5 and 10.1.2 for an explanation of why [β], [ð], [ʝ], and [ɣ] are classed here as fricatives rather than as approximants.

ones (2a) or those occurring immediately after a stressed nucleus (2b), a fact which would correlate with part of the affricate being moraic, in coda position (Recasens 1993; Jiménez 1996a; Bonet & Lloret 1998: 176–8).

(2) (a) *dotze* [ˈdod͡d͡zə] 'twelve'
atzar [əd͡dzar] 'chance'
pitja [ˈpid͡dʒə] 'press.3SG.PRS.IND'
adjectiu [əd͡dʒəkˈtiw] 'adjective'
(b) *fletxa* [ˈflet͡tʃə] 'arrow'
(c) *lletsó* [ʎəˈt͡so] 'dandelion' (*llet* 'milk' + ?)
potser [puˈt͡se] 'maybe' (*pot* 'may' + *ser* 'be.INF')
(d) *gats* [ˈgat͡s] 'cats' /gat+z/
pots [ˈpɔt͡s] 'can.2SG.PRS.IND' /pɔd+z/

Moreover, the frequency of occurrence in morphemes of the invariant affricates is not high compared to other consonants, a fact which is in accord with their being interpreted as phoneme sequences. Lexical intervocalic [t͡s] is very rare anyway (2c). It arises only in what are historically compounds, though probably not so synchronically. These particular words (2c) are realized [ʎəkˈso], [puˈðe], by many speakers. The non-continuant and continuant parts of word-final [t͡s] always correspond to separate morphemes (2d). The voiced alveolar affricate counterpart [d͡z] is never word-final—more strictly, intervocalic [d͡z] does not alternate morphologically with word-final [t͡s], since coda-devoicing would be expected—a fact which again argues strongly in favour of its interpretation as /d.z/; after a tautomorphemic cluster of increasing sonority a vowel is required to construct a well-formed syllable (§3.1 and Chapter 8). Furthermore, unlike most unit phonemes, affricates do not occur after other consonants, or rather, affricates are not in contrast with fricatives in that position. *Clenxa* 'hair parting' may be pronounced either [ˈklɛnʲt͡ʃə] or [ˈklɛnʲʃə], but [nʲt͡ʃ] and [nʲʃ] are not in contrast in any variety; see §2.1.3 below.

Recasens proposes (1993: 161) that cases of [d͡ʒ] ([t͡ʃ] by word-final devoicing) which alternate morphologically with [ʒ], such as *boig* [ˈbɔt͡ʃ] 'crazy.M' – *boja* [ˈbɔʒə] 'crazy.F', should be regarded as unit phonemes, in contrast to the cases in which final [d͡ʒ] ([t͡ʃ]) alternates with [d͡dʒ] as in *mig* – *mitja* 'half', or the cases in which [d͡ʒ] does not alternate as in *viatge* [biˈad͡dʒə] 'journey'; in these we would have phoneme sequences of stop plus fricative.

Reflecting the general markedness of voiced geminate obstruents, there is a tendency to realize [d͡dʒ], and to a lesser extent [d͡dz], as voiceless, in examples like those of (2a). This phenomenon is most often found in continental Catalan from Barcelona northwards (Recasens 1991b: 215–16), though it is by no means general geographically and is subject to social variation where it does occur. Variation in the realization of intervocalic /dʒ/ was one of the phenomena investigated by Pons in Barberà del Vallès in 1983 (Pons 1992). In contrast with what Recasens and others report, she observed a voiced *geminate* [d͡dʒ] only very rarely. But [t͡ʃ] was a common realization. Among her group of twenty-eight informants of different ages (four born per decade between 1900 and 1970) she found some 66 per cent of [t͡ʃ] realizations, in words like *petjada* 'footprint', *lletja* 'ugly.F', *desitjos* 'desires', where /dʒ/ occurs after a vowel which is stressed

in the lexical root. (She found 0 per cent [t͡ʃ] for /dʒ/ which is always pretonic as in *adjectiu* 'adjective' or *suggerir* 'suggest'.) Though all speakers in the sample used both variants, [d͡ʒ] and [t͡ʃ], they differed greatly in the proportions. The variance was also large in each decade cohort, and she concluded there was no evidence of a change in progress. Nor was either variant stigmatized in her evaluation study. In Barberà, voiceless realizations were not found in significant numbers for /dz/.

Mier (1986) investigated voiced and voiceless realizations of intervocalic /dʒ/ in Barcelona using Labovian methodology. She likewise found that individual speakers had marked preferences for one or the other variant, with voiceless realizations being commoner among those born after 1944. Contrary to her expectations, she found voiceless realizations were more common in pretonic position (including in morphemes without stress alternation, e.g. *pitjor* 'worse'). Mier's evidence too did not clearly support the view that a change is in progress involving this phenomenon.

2.1.3 *Fricatives*

2.1.3.1 *Labiodental fricative*

Among the fricatives, the voiced labiodental /v/ (corresponding to orthographic *v* as in *viva* 'alive.F') is restricted to certain dialects, namely, Balearic and southern Valencian, where it is generally quite firm, and varieties of northern Valencian and in southern Catalonia, where it is recessive. In the remaining varieties /v/ has merged with [b] ∼ [β] (Recasens 1991b: 195).

2.1.3.2 *Alveolo-palatal fricatives*

As mentioned in §2.1.2, the voiced lamino-alveolo-palatal fricative /ʒ/ is absent from Valencian,[3] in most areas of which we find only /d͡ʒ/, which is consistently realized [d͡ʒ], corresponding to both [ʒ] and [dd͡ʒ] ∼ [d͡ʒ] elsewhere (3a,b). With the exception of the cases discussed in §2.1.3.4, the voiceless alveolo-palatal fricative /ʃ/ (3c) is contrastive in all varieties.

			E. Catalonia	Valencian
(3)	(a)			
		girar 'turn$_V$'	ʒiˈɾa	d͡ʒiˈɾaɾ
		enveja 'envy'	əmˈbɛʒə	eŋˈved͡ʒa
	(b)	*imatge* 'image'	iˈmad͡ʒə	iˈmad͡ʒe
	(c)	*créixer* 'grow.INF'	ˈkɾeʃə	ˈkɾejʃeɾ
		dibuix 'drawing'	diˈβuʃ	diˈβujʃ

In northern Valencian the single voiced alveolo-palatal phoneme is realized as [d͡ʒ] in initial or post-consonantal position and as [ʒ] or [jʒ] intervocalically (Gimeno Betí 1994: 29–59). This distribution recalls the distribution of other voiced stop–fricative phoneme sets, that is, with fricative realizations in intervocalic onsets and stops elsewhere; see §10.1.2. Even in those dialects where [ʒ] and [d͡ʒ] contrast intervocalically (see *enveja* vs. *imatge* in (3a,b)), there is no

[3] Except in the region of La Marina Alta, where there is a contrast of intervocalic /ʒ/ and /d͡ʒ/ with essentially the same lexical distribution as is found in northern Catalonia (Colomina 1985).

contrast in final position, where only the affricate is found. There are consequent morphophonemic alternations, as in the case of *passejar* [pəsəˈʒa] 'take a walk' – *passeig* [pəˈsɛt͡ʃ] 'walk$_N$' (with word-final devoicing; see Chapter 5). The voiceless counterpart of [ʒ], i.e. [ʃ], does occur in word-final position, as in (3c) and *calaix* [kəˈlaʃ] 'drawer', *boix* [ˈboʃ] 'box tree'. In no variety are alveolo-palatal fricatives and affricates in contrast in initial or post-consonantal position. In western Catalan, which includes Valencian, and in the Tarragona region, affricates are favoured in this initial or post-consonantal context, as in the case of *girar* in (3) or the examples in (4a), though the fricative realization is not foreign to north-west Catalan (Recasens 1991b: 283). In the case of the initial voiceless alveolo-palatal, in some words, many such dialects prefer a lexical variant with an initial vowel and an intervocalic fricative (4b) (Recasens 1991b: 284; Borràs 1992: 325).

(4)		W. Catalan	E. Catalan
(a)	*xinxa* 'bedbug'	t͡ʃinʲt͡ʃa	ˈʃinʲʃə ~ ˈt͡ʃinʲt͡ʃə
	xalet 'bungalow'	t͡ʃaˈlet	ʃəˈlɛt ~ t͡ʃəˈlɛt
	jove 'young'	ˈd͡ʒove	ˈʒoβə ~ ˈd͡ʒoβə
	taronja 'orange'	taˈrɔnʲd͡ʒa	təˈrɔnʲʒə ~ təˈrɔnʲd͡ʒə
(b)	*xarop* 'syrup'	ajʃaˈrɔp	ʃəˈrɔp ~ t͡ʃəˈrɔp
	Xàtiva (toponym)	ajˈʃativa	ˈʃatiβə

In eastern Catalan a fricative is regarded as 'standard' in these word-initial and post-consonantal non-contrastive contexts, though in fact, in the Barcelona region at least and probably more widely, fricative and affricate realizations are in something like free variation. In her 1983 sociolinguistic study in Barberà del Vallès (20 km from Barcelona), Pons (1992: 324) observed slightly under 50 per cent affricate realizations for initial /ʃ/ (in words spoken in isolation). She found a lower index of affrication for initial /ʒ/, with a mean of 35 per cent, with some speakers avoiding initial [d͡ʒ] altogether. There was evidence of a historical trend, but not a linear one. (Affrication increased notably in the four decade cohorts from the 1900s to the 1940s, but fell back subsequently.) In listening tests, the majority of informants could not distinguish between the two variants [ˈʒoβə] and [ˈd͡ʒoβə] for *jove*, and even those who could showed no preference for one or the other.

In her survey of forty-three Barcelona informants, for the voiceless alveolo-palatals, Mier (1986) observed 86 per cent word-initial affricate [t͡ʃ] realizations of /ʃ/ in interview style. For these word-initial cases of /ʃ/, the preceding segment was of no influence. For the voiced counterpart though, the preceding context did make a difference, with 73 per cent [d͡ʒ] word-initially after [n] (as in *un jove* 'a young man'), 66 per cent [d͡ʒ] in utterance-initial position before a stressed vowel (*jove*), 44 per cent [d͡ʒ] in utterance initial position before an unstressed vowel (as in *gener* 'January'), and 5 per cent word-initially after a vowel (as in *la gent* 'people'). Mier is imprecise about numbers for medial post-consonantal position, though it is clear that affricate variants were commoner for the voiceless alveolo-palatal (*marxar* 'leave') than for the voiced one (*targeta* 'card'). She interpreted her evidence (which had many fewer examples than Pons's study) as hinting at a

trend towards more affricate realizations, which she argued is subject to lexical diffusion in the case of [t͡ʃ] ∼ [ʃ]. From a typological perspective, in the absence of a phonemic contrast, an affricate might well be preferred as a 'stronger' onset in a prominent word-initial position. (There is some evidence too for the general marked status of fricative [ʒ] vis-à-vis other fricatives; compare the absence of /ʒ/ in Italian (beside /ʃ/, /dʒ/, etc.), the marginal status of /ʒ/ in English (also alongside /ʃ/, /dʒ/, etc.), and so on.)

2.1.3.3 *Relation between alveolo-palatal fricatives and affricates*

I review in this section the question of how to interpret variation between an affricate and a fricative of the same place of articulation, where these are not in contrast, and of [ʒ] versus [dd͡ʒ] ∼ [d͡ʒ] where they are in contrast. To do so it will be necessary to consider some aspects of syllabification, which will be taken up in general in §3.1. The voiced and voiceless pairs are taken separately because, as previously mentioned, their situations are not exactly parallel.

In Valencian, except for La Marina Alta, the situation is clear (see §2.1.3.2): we find either invariant [d͡ʒ] in all positions or, in northern Valencia, [ʒ] or [jʒ] in medial intervocalic position corresponding to [d͡ʒ] in other positions. It is quite plausible to view [ʒ] or [jʒ] in these northern Valencian varieties as lenition variants of /d͡ʒ/, though lenition of /d͡ʒ/ seems to be less general than that of the remaining voiced stops (§10.1.2). As far as other varieties are concerned, I compare three alternative interpretations (A, B, C) with representative contexts in (5).

(5)	Examples	Realization (ignoring word-final devoicing)	Interpretation A Input /d͡ʒ/	Interpretation B Input /ʒ/	Interpretation C Input both /d͡ʒ/ and /ʒ/, underspecified /J/
(a)	*gent* 'people' *menjar* 'eat'	[d͡ʒ] ∼ [ʒ]	#d͡ʒV $C_{[+sonorant]}$d͡ʒV	#ʒV $C_{[+sonorant]}$ʒV	#JV $C_{[+sonorant]}$JV
(b)	*roja* 'red.F' *rajar* 'flow$_V$'	[ʒ]	Vd͡ʒV	VʒV	VʒV
(c)	*corretja* 'belt' *pitjor* 'worse'	[dd͡ʒ] ∼ [tt͡ʃ] ∼ [d͡ʒ] ∼ [t͡ʃ]	Vdd͡ʒV = Vd͡ʒd͡ʒV	Vd ʒV	Vd͡ʒV
(d)	*roig* 'red.M' (cf. *roja*.F)	[d͡ʒ]	Vd͡ʒ#	Vʒ#	Vʒ#
(e)	*desig* 'desire$_N$' (cf. *desitjar* 'desire$_V$')	[d͡ʒ]	?	?	Vd͡ʒ#
(f)	*viatge* 'journey' (cf. *viatjar* 'travel$_V$')	[dd͡ʒə] ∼ [tt͡ʃə] ∼ [d͡ʒə] ∼ [t͡ʃə]	Vd͡ʒd͡ʒ# =Vdd͡ʒ#	Vd ʒ#	Vd͡ʒV# or Vdd͡ʒ#
(g)	*conserge* 'doorman'	[d͡ʒə] ∼ [ʒə]	$C_{[+sonorant]}$d͡ʒ#	$C_{[+sonorant]}$ʒ#	$C_{[+sonorant]}$ʒ# or $C_{[+sonorant]}$JV#

Interpretation A in (5) has one underlying phoneme /d͡ʒ/, which is subject to lenition (→ [ʒ]), and may occur geminate /d͡ʒd͡ʒ/ (or equivalently /d͡ʒd͡ʒ/ or /d͡ʒ/). Interpretation B has one underlying /ʒ/, which is subject to fortition (→ [d͡ʒ]) in the complementary contexts to A's lenition; /ʒ/ may be preceded by an obstruent stop, so the sequence /d ʒ/ is realized as [d͡ʒd͡ʒ] or [d͡ʒ]. Interpretation C has two phonemes, an affricate /d͡ʒ/ and a fricative /ʒ/; the contrast between them is neutralized in some environments. Here /J/ symbolizes a voiced alveolo-palatal obstruent unspecified for [continuant]. Consider now what would need to be said in accounting for the outputs in each of the contexts illustrated in (5).

Taking interpretation A with input /d͡ʒ/ first, the (variable) output [ʒ] in context (5a) looks like lenition, a phenomenon shared with other voiced obstruent stops in Catalan (§10.1.2). Variation between [d͡ʒ] and [ʒ] in initial and post-consonantal position does not show quite the same proportions as are found with the other stops, though. In comparison with the other stop–fricative pairs, the (optional) presence of [ʒ] after nasals, in particular, may seem a bit surprising: since nasals are stops, and are also homorganic, one might expect an oral stop to be favoured. It could be, however, that [nʲd͡ʒ] lacks an open stop release phase (since the stop is released into the fricative), and thus that the presence of an oral stop, even if intended by a speaker, is hard to detect. It is also the case that there is no surface distinction in Catalan between /nt+s/ and /n+s/, or between /nd+z/ and /n+z/, even though these sequences are independently justifiable on morphological grounds. In the intervocalic context (5b) regular lenition of /d͡ʒ/ to [ʒ] is entirely consistent with categorical lenition of the other voiced stops in this position. In context (5c) the realizations [d͡ʒd͡ʒ], [t͡ʃt͡ʃ], [d͡ʒ], [t͡ʃ] are in line with what is to be expected from an input affricate with a mora attached (i.e. a geminate—the choice of linear IPA representation of the geminate is of no consequence here). The voiceless realizations reflect the general difficulty of maintaining enough intra-oral pressure to sustain voicing in a geminate voiced stop (though voice is contrastive in this context, so lexical information is lost). The non-geminate [d͡ʒ] reflects the fact that output length is not in fact necessary to maintain the lexical contrast between /dd͡ʒ/ and /d͡ʒ/ in a context where the latter is realized as a fricative; realizing /dd͡ʒ/ as [d͡ʒ] involves a venial violation of MAXμ. The short realizations, voiced or voiceless, might be expected to occur especially in the speech of speakers who highly rank *GEMINATE, a ranking which is certainly known to occur in (non-standard) Catalan, though whether short realizations of /dd͡ʒ/ in fact correlate with short realizations of /mm/, /nn/, /ll/, /ʎʎ/ is not known. However, the input Vdd͡ʒV or Vd͡ʒd͡ʒV, as a sequence of voiced stops, is somewhat unusual, though not quite unparalleled. Intervocalic geminate /dd/ and /bb/ are attested (see §2.1.9), as are /bd/, /db/, and /gd/, though the non-affricate voiced stop clusters are almost all in the Greco-Latin learned vocabulary.

The realization (5d) of /d͡ʒ/ as a stop in final position after a vowel calls for no comment in itself; stops are regularly maintained in codas. But the contrast with contexts (5e) and (5f) is trickier. Pattern (5d), *roig – roja*, consists of a word-final affricate that alternates with a fricative intervocalically. Pattern (5e), *desig – desitjar*, consists of a word-final affricate alternating with a (geminate) affricate intervocalically, while (5f), *viatge – viatjar*, illustrates the pattern in which a morpheme-final stop cluster provokes final-vowel 'epenthesis':[4] the output of the stop cluster is always an intervocalic (geminate) affricate (see Chapter 8). If one wishes to maintain a unique input for the root morphemes of the (5e) type, one is obliged to invoke some kind of exceptionality. If we say *desig* has final /d͡ʒ/, the absence of lenition in intervocalic position is exceptional (cf. type (5d)). If we say *desig* has final /dd͡ʒ/ (or /d͡ʒd͡ʒ/), on the basis of the intervocalic variant of the morpheme (*desitjar*), the absence of epenthesis is exceptional (cf. (5f); **desitge* would be expected). Neither of these solutions is appealing, but we have only two input structures, /d͡ʒ/ and /dd͡ʒ/, with which to account for three patterns of morphophonemic alternation (5d, e, f). The most likely way out would be to treat the *desig – desitjar* pattern of alternation as reflecting stem allomorphy: /dezid͡ʒ/ in word-final position and /dezidd͡ʒ/ elsewhere.[5] As will be seen (§10.2–10.3), this is not the only case in which a root-final consonant alternation has become opaque.

Contexts (5f), *viatge*, and (5g), *conserge*, involve the realization of /d͡ʒ/ when morpheme-final, after a (homorganic) oral stop and after a sonorant respectively. What we observe in both contexts is vowel 'epenthesis'. In context (5f), epenthesis is entirely predictable after a stop+stop cluster, since the cluster /dd͡ʒ/ violates the SONORITY SEQUENCE constraint on well-formed syllables (§3.1), and epenthesis repairs this disharmony. The epenthesis is parallel to what is seen in *docte* 'learned', *apte* 'suitable', and so on. But with input sequences of sonorant+voiced stop (5g), epenthesis is surprising. Generally in Catalan a word-final sequence of sonorant+voiced stop is well-formed (subject

[4] I maintain the familiar term 'epenthesis' for situations such as this where a segment and zero are not in contrast. However, the OT strategy of Lexicon Optimization (Prince & Smolensky 1993: 192; Kager 1999: 32–4) requires that underlying forms should match surface forms in the absence of positive evidence to the contrary.

[5] The reason for preferring allomorphy in type (5e), rather than in type (5d), comes down to the number of morphemes involved. The following words are the ones of the (5e) type: *mig* 'half', *lleig* 'ugly', *desig* 'desire', *rebuig* 'rejection', *enuig* 'annoyance', *estoig* 'case', *trepig* 'tread', *safareig* 'washtub', *llebeig* 'south-west wind', and *bolig* 'sweep net', together with whichever other masculine nouns in [-t͡ʃ] speakers may prefer a plural in [-d͡ʒus] for (e.g. *passeig* 'walk'). The list is shorter in areas of northern Catalonia and in La Marina Alta (Valencia), where items like *mig*, *lleig* are in the (5d) class. However the list is longer in Balearic Catalan, where the verb-forming suffix *-ejar* is pronounced with [dd͡ʒ], and corresponds to *-eig* [-d͡ʒ] in associated nominals.

to word-final obstruent devoicing, of course). Some examples are given in (6). The orthography reflects the input.

(6) *corb* 'raven' *tomb* 'turn'
sord 'deaf' *normand* 'Norman'
amarg 'bitter' *fang* 'mud'

From this perspective one should expect, for example, **conserj* /konsɛRd͡ʒ/ [kunˈsɛrt͡ʃ] to be well-formed. There are no such words. And since comparable sequences with a voiceless fricative/affricate do exist, though they are rare (*guerx* 'cross-eyed', *ponx* 'punch'), the absence of **conserj*, etc. would have to be regarded as fortuitous.

Consider now the case for interpretation B with input /ʒ/. In context (5a) one would have to speak of (variable) fortition in strong positions, i.e. word-initially and in post-consonantal onsets, to deal with the [d͡ʒ] variant. This is not implausible, though it would make a more convincing story if a comparable account could be given for the other stop–fricative pairs. Given that /ʒ/ is a sibilant, one would have to explain why the other voiced sibilant /z/ is never subject to fortition in the same environment. No doubt fortition is more marked than lenition, other things being equal, which would count against the B analysis in comparison with A. In context (5b), *roja*, in interpretation B the fricative output exactly matches the input. In context (5c), *corretja*, the account goes much as in interpretation A, except that the account of 'gemination' is marginally more complex. At first sight input Vdʒ̍V could be syllabified V.dʒ̍V. But in Catalan there is not sufficient Minimum Sonority Distance between two obstruents to make a well-formed onset (see §3.1). The syllabification Vd.ʒ̍V, however, violates the SYLLABLE CONTACT law; the sonority of the coda needs to be equal to or greater than the sonority of the onset (§3.1). So Vd.d͡ʒ̍V is more harmonic than either of the alternatives so far considered. Or we arrive at the same outcome in (5c) if the fortition process observed in (5a) is understood to apply equally after coda stops. The underlying sequence of stop+voiced sibilant /dʒ/ is paralleled not only by /-dz-/ (§2.1.2) but also by /-gz-/ and /-bʒ-/. Context (5d), *roig*, is somewhat problematic for the B account, rather as (5a) is. We have coda fortition, which is otherwise characteristic of non-strident obstruents ([β]→ [b], [ð]→ [d], [ɣ]→ [g]), not of strident ones. Strident /z/ remains a fricative in a coda, while /v/, in those dialects that have it, weakens to [w], as in *nova* [ˈnɔvə] 'new.F' – *nou* [ˈnɔw] 'new.M'. In analysis B the difficulties with the interpretation of pattern (5e), *desig*, are analogous to those that occur with analysis A, since there are only two contrasting input structures, /ʒ/ and /dʒ/, with which to account for three morphophonemic patterns. Input /deziʒ/ for *desig* – *desitjar* would involve regular word-final fortition, as in (5d), but entirely anomalous exceptional intervocalic fortition in *desitjar*. Input /dezidʒ/ would have the same problem as the counterpart analysis in approach A, i.e. unmotivated lack of epenthesis after a string violating the SONORITY SEQUENCE constraint. Allomorphy would require /deziʒ/ for *desig* but /dezidʒ/ for *desitjar*. Context (5f), *viatge*, however, would

illustrate expected epenthesis after a sequence of segments /dʒ/ violating the SONORITY SEQUENCE constraint. In context (5g), *conserge*, the epenthesis outcome has some parallels, involving a stem-final sequence of sonorant+voiced sibilant, comparable to what is observed in *catorze* 'fourteen' or *bronze* 'bronze', though why underlying voicing in particular should provoke epenthesis is obscure.

To sum up so far: input interpretation A appears superior for contexts (5a) and (5d); interpretation B looks marginally superior for context (5g). They are probably equally suitable in context (5c) and (5f). Neither account provides a neat solution for (5e), though if 'exceptionality' is employed, interpretation A has a better story; if allomorphy is involved, the two accounts are equal.

Interpretation C has both a fricative and an affricate as unit phonemes; the contrast between them is neutralized in many environments. This account is parallel in several respects to an analysis of /r/ and /ɾ/ as phonemes that happen to contrast only in the intervocalic environment (§2.1.6), though major phonetic difference between the rhotics and the alveolo-palatal sibilants means that any parallels are hard to exploit in defending an interpretation. In context (5a), *gent*, *menjar*, we would observe neutralization of contrast between /d͡ʒ/ and /ʒ/. There is no obvious reason for neutralization in the 'strong' word-initial environment, at least, though after a homorganic sonorant the contrastive presence or absence of a stop element is hard to detect (cf. English *mince* vs. *mints*). The contrast manifested between contexts (5b), *roja*, and (5c), *corretja*, is simply dealt with as a contrast between two different unit phonemes. However the variable geminate and voiceless realizations of an intervocalic single affricate /d͡ʒ/ defy a natural account. The morphophonemic contrast between patterns (5d), *roig*, and (5e), *desig*, receives a superficially simple account: the two different input phonemes /ʒ/ and /d͡ʒ/ are involved, which neutralize as [d͡ʒ] in word-final position; final /ʒ/ undergoes fortition, as in interpretation B, but now we cannot so easily exploit parallels with the fortition of non-strident obstruents in codas ([β]→ [b], [ð]→ [d], [ɣ]→ [g]), since with the non-stridents the alternations are allophonic. In context (5f), *viatge*, interpretation C might offer two alternatives: either /Vd͡ʒV#/ or /Vdd͡ʒ#/. The first of these makes (5f) identical with (5c), and the 'epenthetic' aspect of the final vowel in *viatge* is ignored. The alternative Vdd͡ʒ# makes sense of the epenthesis, to repair the SONORITY SEQUENCE violation but now underlying gemination is only contingently associated with surface gemination. They match in type (5f), but in types (5c) and (5e) surface gemination corresponds with an underlying single stop. (This seems to be essentially the analysis that Recasens (1993: 161) has in mind.) Generalizing the /dd͡ʒ/ pattern to (5c) and (5e) types would make interpretation C a superficial variant of interpretation A, with the disadvantage of implausibly allowing a geminate stop in a context where the corresponding simple stop was excluded. Alternative solutions are also available to interpretation C for context (5g): either $C_{[+\text{sonorant}]}$ʒ#, which has the benefits of this input in interpretation B in making epenthesis follow naturally as a repair to the Minimum Sonority Distance violation, or $C_{[+\text{sonorant}]}$JV#, which

draws on the surface identity with the post-sonorant prevocalic context of (5a). The strength of interpretation C lies in its account of the contrast between (5d), *roig*, and (5e), *desig*, as corresponding to an underlying difference of contrasting units. However, it then hasn't a coherent story to tell for type (5f), *viatge*.

The possible approaches to the contrastive status of [ʃ] and [t͡ʃ] are, naturally, parallel in certain respects to those for [ʒ] and [d͡ʒ], but there are some important differences, both of fact and of plausible interpretation. In the first place, in the case of the voiceless alveolo-palatals, differently from the voiced ones, Valencian, just like other dialects, has a clear contrast between [ʃ] and [t͡ʃ] in intervocalic position (*aixa* [ˈaʃa] ∼ [ˈajʃa] 'adze', *atxa* [ˈat͡ʃa] 'torch'). Secondly, in word-final position there is a straightforward contrast between [ʃ] and [t͡ʃ], each of which also appears unaltered in related intervocalic forms (7).

(7) *baix* [baʃ] 'low' *baixar* [bəˈʃa] 'lower$_V$'
despatx [dəsˈpat͡ʃ] 'office' *despatxos* [dəsˈpat͡ʃus] 'offices'

It is only in word-initial and post-consonantal positions that the contrast between a voiceless alveolo-palatal affricate and a fricative is absent. Their distribution has been described in §2.1.3.2. Some Valencian varieties, at least, allow a contrast between initial [ʃ] and [t͡ʃ] as well, albeit somewhat marginally, for example among Valencian toponyms such as those in (8) (Borràs 1992; Sanchis Guarner 1993: 92–4).

(8) *Xixona* [ʃiˈʃona] *Xiva* [t͡ʃiva]
Xeresa [ʃeˈɾeza] *Xert* [ˈtʃɛɾt]
Xeraco [ʃeˈɾako] *Xivert* [t͡ʃiˈvɛɾt]

For such Valencian varieties, at least, a phonemic contrast /t͡ʃ/ versus /ʃ/ looks the most plausible solution. For the others, an interpretation like type A in (5) does not look well-founded. Input /t͡ʃ/ would correspond to variable (initial and post-consonantal) [t͡ʃ] ∼ [ʃ], and to invariant medial and final [ʃ], with unparalleled 'lenition' of a voiceless stop, and, word-finally, lenition in an unparalleled environment. Surface [t͡ʃ] would derive from a sequence /tt͡ʃ/, which would be helpful for medial geminate realizations but would be problematic in final position where epenthesis would be expected after a stop+stop sequence. Interpretation B for the voiceless alveolo-palatals would have underlying /ʃ/, subject to variable fortition in initial and post-consonantal onsets, and able to occur after /t/ to derive intervocalic and word-final (invariant) affricates, potentially geminate in so far as [t.t͡ʃ] realized a better syllable contact and Minimum Sonority Distance than [.tʃ] (or [t.ʃ]). But in morpheme-final word-final position we might expect /t͡ʃ/ to demand an epenthetic vowel, which examples like *despatx* (7) show not to be necessary.[6] Epenthesis is the norm in the case of variable [ʃ] ∼ [t͡ʃ] after

[6] There are some 20 nouns of the *despatx* type, showing no epenthesis after final [t͡ʃ]; 5 nouns have an 'epenthetic' vowel after [t͡ʃ] (*apatxe* 'apache', *fantotxe* 'nonentity', *fetitxe* 'fetish', *cotxe* 'car', and *reprotxe* 'reproach'). The masculine gender suffix /+o/ is markedly common after [t͡ʃ]—some 20

a sonorant in morpheme-final position, as in *porxo* 'portico', *ganxo* 'hook'. All the examples in fact have masculine /+o/. The only exceptions to epenthesis in this context appear to be *guerx* 'cross-eyed' which has a preferred variant *guerxo*, the Anglicism *ponx* 'punch', and the Valencian toponym *Elx* (and Valencian, I argue, has /t͡ʃ/ in any case).

In an interpretation of type C, there would be two phonemes /t͡ʃ/ and /ʃ/ (as there need to be in the Valencian varieties mentioned above in (8)). They would be in contrast only after a vowel; neutralization initially and after a consonant would permit 'free' variation between affricate and fricative. This interpretation is consistent with Mier's (1986), mentioned above (§2.1.3.2), according to which initial [t͡ʃ] is spreading by lexical diffusion. In principle, interpretation C allows the unit /t͡ʃ/ and the sequences /tʃ/ or /t t͡ʃ/ to be in contrast. Either of the sequences might underlie an intervocalic geminate realization. As mentioned, geminate realizations are reported; lexical contrast between a geminate and a single affricate has not been observed, though. Correspondingly, in morpheme-final position the unit /t͡ʃ/ should provoke no epenthesis, while the sequence /tʃ/ or /tt͡ʃ/ would require to be syllabified with a following vowel. I know of no independent evidence for attributing the difference between *despatx* and *totxo* to an input contrast of this type (see note 4 above).

In the case of the alveolo-palatal stridents, then, evidence does not lead overwhelmingly to one analytical solution. Nor does the evidence point strongly towards a matching solution for the voiced and the voiceless sets. Forced to choose, as a linguist, I might well prefer interpretation A for the voiced alveolo-palatals (input /d͡ʒ/ with lenition, sequence /dd͡ʒ/ leading to gemination, allomorphy for the *desig* – *desitjar* alternations) and interpretation C for the voiceless alveolo-palatals (input contrast of /ʃ/ and /t͡ʃ/, with neutralization non-prevocalically). But who but a linguist is forced to make such a choice? There is little evidence that speakers are forced to do so, or indeed, that if they do make such a choice they agree among themselves on the outcome. In fact, variation in realization in prevocalic position—affricate or fricative initially or after a sonorant, single or geminate affricate intervocalically—may well reflect different, and even inconsistent, phonological interpretations by speakers. Different, even inconsistent, analyses may also underlie the patterns of morphophonemic alternation illustrated in (5b–d), (5e), and (5f): affricate–fricative alternation in *roig* – *roja*; affricate–(geminate) affricate alternation without final epenthesis in *desig* – *desitjar*; and (geminate) affricate–(geminate) affricate alternation with epenthesis in *viatge* – *viatjar*. As mentioned above, different regional dialects clearly make some different choices here; most Valencian merges the first two, with different results in the centre and south ([d͡ʒ] only) from those in the north ([d͡ʒ]—[ʒ] only). Optimality Theory allows the linguist more directly to model the behaviour of

nouns such as *cucurutxo* '(ice cream) cornet' and *totxo* 'brick'. While this fact is consistent with a biphonemic interpretation of [t͡ʃ] as a cluster /t͡ʃ/ of increasing sonority, which therefore requires a following vowel, it surely does not impose such an interpretation.

speakers in situations like these, since in OT the nature of the input representation is of no particular importance. The task is to identify the constraints, and their ranking, that govern well-formed outputs. The 'cost' is that one cannot assume a basic underlying representation that surfaces by default (through Faithfulness constraints). The logic of Lexicon Optimization and 'Richness of the Base' (Smolensky 1996; McCarthy 2003a) is that *all* occurring allophones need an explicit account. Thus when [d͡ʒ] and [ʒ] alternate, the constraints need to be identified that favour the correct output whichever occurs in the input representation. And when [d͡ʒ] and [ʒ] are in contrast, the account ought not to depend on details of input representation, provided the input contrast is 'notated' in some way that is reflected naturally in the output.

2.1.3.4 *Interpretations of* [ʃ]

In some varieties of northern Valencia and southern Catalonia /ʃ/ has merged with realizations of /s/ after a high front vocoid (Rafel 1978): for example, *terrissa* [teˈriʃa] 'pottery', *pixa* [ˈpiʃa] 'piss.3SG.PRS.IND', *insistisc* [insiʃˈtiʃk] 'insist.1SG.PRS.IND', *caixa* [ˈkajʃa] 'box', *deixa* [ˈdejʃa] 'leave.3SG.PRS.IND', *cuixa* [ˈkujʃa] 'thigh' (these last three are e.g. [ˈkaʃə], [ˈdeʃə], [ˈkuʃə] in other varieties). In these varieties [ʃ] is not found after other vocoids, and merges with /t͡ʃ/ after consonants (*punxa* [ˈpunt͡ʃa] 'point'). Thus [ʃ] is an allophone of /s/. Jiménez (1996b) puts forward the case that [ʃ] is an allophone of /s/, not only in the varieties just mentioned but in all varieties in which *caixa* 'box' and *baix* 'low', etc. are pronounced [ˈkajʃa] and [ˈbajʃ] rather than [ˈkaʃa] and [ˈbaʃ], i.e. in most western Catalan and the southern part of continental eastern Catalan. This account runs into problems, some of which Jiménez confronts, such as the fact that most such dialects retain some contrast between /s/ and /ʃ/ in word-initial position. Relevant examples of initial /ʃ/ have been given above at (8). Another difficulty, which I think is insuperable, is the fact that such dialects, unlike the ones in northern Valencia and southern Catalonia mentioned at the beginning of this section, retain a contrast between /s/ and /ʃ/ after /i/. Some examples illustrating this are given in (9), in Valencian pronunciation.

(9)	/s/	/ʃ/
	Vicent [viˈsent] (given name)	*ixent* [iˈʃent] 'rising'
	Benissa [beˈnisa] (toponym)	*pixa* [ˈpiʃa] 'piss.3SG.PRS.IND'
	feliç [feˈlis] 'happy'	*Flix* [ˈfliʃ] (toponym)
	morisca [moˈriska] 'Moorish.F'	*agraïxca* [aɣɾaˈiʃka]'thank.3SG.PRS.SUBJ'
	mestís [mesˈtis] 'cross-breed'	*tix* [ˈtiʃ] 'weave.3SG.PRS.IND'

2.1.3.5 *Dorso-palatal fricative* [ʝ]

The dorso-palatal fricative [ʝ] is found, in complementary distribution with [ɟ], only in Majorcan varieties that have dorso-palatals rather than the velars found in most dialects.

Majorcan

guerra [ˈɟɛrə] 'war'	*sa guerra* [səˈɟɛrə] 'the war'
crec [ˈkɾəc] 'believe.1SG.PRS.IND'	*cregui* [ˈkɾəʝi] 'believe.1SG.PRS.SUBJ'

In recent work on Catalan phonology (such as Recasens 1991b; 1993; Bonet & Lloret 1998) the bilabial, dental, and velar continuants [β], [ð], [ɣ] are referred to as approximants rather than as fricatives. Recasens's reasons for this are based on acoustic evidence; these sounds lack the turbulence usually regarded as characteristic of fricatives. However, these sounds behave in Catalan as obstruents. They alternate morphophonemically with the voiced and voiceless plosives of corresponding place, and form syllable onset clusters with liquids [.βɾ], [.ðɾ], [.ɣɾ], [.βl], [.ɣl] in a way consistent with their being of lower sonority than liquids. I believe it is more appropriate to regard them as non-strident fricatives ('spirants' in older terminology), which is also typologically consistent with their being the product of lenition processes (see §10.1.3 and Kirchner 1998). Following Recasens's interpretation, the Majorcan dorso-palatal fricative [ʝ] is indistinguishable from the dorso-palatal approximant [j], whereas we need to reflect contrasts such as seen in Majorcan *pagam* [pəˈʝam] 'pay.1PL.PRS.IND' vs. *treballam* [trəbəˈjam] 'work.1PL.PRS.IND'. Calling spirants 'non-strident fricatives' is not inconsistent with Recasens's description of their low acoustic turbulence.

2.1.3.6 *Varieties without voiced strident phonemes*

Central Valencian, and most Catalan varieties in Aragon, have no voiced strident obstruent phonemes. In these dialects the /z/ and /d͡ʒ/ of other dialects have merged with their voiceless counterparts, /s/ and /t͡ʃ/. This pronunciation is known in Catalan as *apitxat* (or, as some prefer to write it, *apitjat* 'squeezed'). The Valencian version, at least, is regarded as non-standard and is stigmatized. Thus in Valencian *apitxat* we find pronunciations as in (10), with examples selected from (2a) and (5).

(10)

dotze [ˈdot͡se] 'twelve'	*roja* [ˈrot͡ʃa] 'red.F'
atzar [at͡sar] 'chance'	*rajar* [rat͡ʃaɾ] 'flow$_V$'
pitja [ˈpit͡ʃa] 'press.3SG.PRS.IND'	*pitjor* [pit͡ʃoɾ] 'worse'
adjectiu [at͡ʃekˈtiw] 'adjective'	*viatge* [biˈat͡ʃe] 'journey'
gent [ˈtʃent] 'people'	*conserge* [konˈseɾt͡ʃe] 'doorman'
menjar [menʲt͡ʃaɾ] 'eat'	

The same varieties are among those that have merged /v/ with (non-strident) [b] ∼ [β]. The absence of voiced stridents is an areal phenomenon that these varieties share with Spanish, Galician, and Basque; *apitxat* is generally attributed to Spanish contact influence, though it is of some antiquity.

2.1.4 *Nasals*

Of the nasals /m/, /n/, and [ŋ], the velar [ŋ] occurs only before velar obstruents, though the obstruents may be suppressed on the surface, as in *bancs* [baŋs] /bank+s/ 'benches' (§7.2). The apico-dental /n/ participates in the typologically frequent kind of place assimilation to following consonants. Alveolo-palatal /ɲ/ is rare in word-initial position (there are some twenty morphemes displaying it). Most such words are affective in connotation (e.g. *nyeu-nyeu* [ˌɲɛwˈɲɛw] 'hypocritical talk'), but not all (e.g. *nyora* [ˈɲorə] 'pebble').

2.1.5 *Glides (semivowels)*

Though the non-syllabic realization [j] and [w] of high vocoids is often predictable—they occur adjacent to another vocoid when a syllable margin is more harmonic than a hiatus—there are several contexts in which this is not so; this topic is the focus of §§3.2.2–3.2.5.

2.1.6 *Rhotics*

Two rhotic types contrast in Catalan, in intervocalic position only: alveolar trill in, for example, *serra* [ˈsɛrə] 'saw', and alveolar tap in, for example, *cera* [ˈsɛɾə] 'wax'. The contrast is neutralized in all other environments. The description of the non-contrastive distribution and the phonological interpretation of both the contrastive and non-contrastive distribution have occasioned some controversy in the phonological literature. I begin with an account of the sound types and their distribution, following Recasens (1993: 162, 176–8; 1991b: 324–40).

2.1.6.1 *Distribution of rhotic types*

A trill, with from two to four contacts, is found in a syllable onset at the beginning of a root or a lexical prefix (11a), after a heterosyllabic consonant (11b), and between vocoids word-internally (11c). Only in the last of these contexts is a contrast with a tap available.

(11)	(a)	*ros* [ˈros] 'fair'	*enregistrar* [ənrəʒisˈtra] 'record.INF'
		pre-romà [preruˈma] 'pre-Roman'	*refer* [rəˈfe] 're-make'
	(b)	*Enric* [ənˈrik] (given name)	*folro* [ˈfolru] 'lining'
	(c)	*ferro* [ˈfɛru] 'iron'	*Sarrià* [səˈrja] ∼ [səriˈa] (district of Barcelona)

In the contexts of (11) a greater number of tongue-tip contacts is favoured by the presence of an adjacent stressed vowel, or of a preceding plosive (both factors are involved, for example, in *pot riure* [pɔd ˈriwɾə] 'can.3SG.PRS.IND laugh'). A lesser number of contacts is favoured by a preceding fricative or approximant. In north Catalonia, and in the town of Sóller (Majorca), a uvular trill ([ʀ]) or approximant ([ʁ]) can be heard instead of an alveolar trill.

Variation between a short trill (two contacts, here represented [ř]) and a tap is found in coda position before a consonant other than an approximant (or non-strident fricative) (12a), or word-finally after a stressed vowel before a pause (12b). The latter realization may be voiceless ([ɾ̥], [ř̥]).

(12) (a) *forma* [ˈfořmə] ~ [ˈfoɾmə] 'shape'
amarg [əˈmařk] ~ [əˈmaɾk] 'bitter'
autocar parat [əwtuˈkař pəˈɾat] ~ [əwtuˈkaɾ pəˈɾat] 'stationary coach'
mar salada [ˈmař səˈlaðə] ~ [ˈmaɾ səˈlaðə] 'salty sea'
(b) *atur* [əˈtuř] ~ [əˈtuɾ] ([əˈtuř̥] ~ [əˈtuɾ̥]) 'unemployment'

The variation between a short trill and a tap in the contexts of (12) is largely dialectal. The short trill is reported to be typical of continental Catalonia (except in the south-west) and of Minorca; the tap is typical of the remaining regions, i.e. Valencia, south-west Catalonia, Majorca, and Ibiza. Pronouncing dictionaries (such as Bruguera 1990) that represent central Catalan use the symbol [r] for the short trill, and I copy this practice generally in this work in transcriptions representing the relevant varieties. But even in the dialects where the short trill is to be found, a tap may occur after an unstressed vowel, or in fast speech.

In other contexts a tap [ɾ] is found. These contexts are: within words between vocoids (13a), where a contrast with [r] ((10c) above) is available; between vocoids where [ɾ] is morpheme-final (13b); in a coda before an approximant or non-strident fricative (13c); after a consonant in a complex onset (13d); and before a pause, after an unstressed vowel (13e).

(13) (a) *pare* [ˈpaɾə] 'father'
centurió [səntuˈɾjo] 'centurion'
(b) *quedarà* [kəðəˈɾa] 'stay.3SG.FUT'
quedar-hi [kəˈðaɾi] 'to stay there'
mar Egeu [ˈmaɾ əˈʒɛw] 'Aegean sea'
(c) *herba* [ˈeɾβə] 'grass'
sabor delicat [səˈβoɾ ðəliˈkat] 'delicate flavour'
autocar iugoslau [əwtuˈkaɾ juɣuzˈlaw] 'Yugoslav coach'
(d) *prou* [ˈpɾɔw] 'enough'
agradar [əɣɾəˈða] 'to please'
refresc [rəˈfɾɛsk] 'refreshment'
(e) *màrtir* [ˈmařtiɾ] 'martyr'
caràcter [kəˈɾaktəɾ] 'character'

An alveolar approximant [ɹ] is a possible variant of the tap [ɾ] in all post-vocalic contexts, in the varieties that avoid the short trill in (12), i.e. in Valencia, south-west Catalonia, Majorca, and Ibiza. The table in (14), modified from Recasens (1986: 77; and cf. Serra 1996b: 232), summarizes the distribution of rhotics in terms of the number of tongue-tip contacts observed in the various contexts. In (14) manners of articulation among consonants are not distinguished. The symbol V̆ represents an unstressed vowel.

(14)	contexts	V.ɾV V.ɾ#V	.CɾV	V̆ɾ‖	VɾC̆	ˈVɾ̆‖	V.rV	C.rV	C#rV ‖rV
	contacts	1	1	1	1–2	1–2	2	2–3	2–4

2.1.6.2 *Interpretations of rhotic distribution*

In the pre-generative tradition (Alarcos Llorach 1953; Badia i Margarit 1965) the phonemic contrast between tap /ɾ/ and trill /r/ having been established, for example, by minimal pairs such as *serra* vs. *cera* (see §2.1.6), no further explanation was offered of their very restricted distribution, namely contrast in intervocalic position only, with neutralization in favour of either the trill or the tap in all other positions. Since then, four different proposals have been put forward to make sense of the unbalanced distribution. The 'classic' analysis within generative phonology (Wheeler 1975/9: 191–4; 1987; Bonet & Lloret 1998: 83–93) interpreted the intervocalic trill [r] in *serra* as realizing geminate /rɾ/ (or /RR/). Thus there was one phoneme (or underlying segment) /ɾ/, which, like other sonorants, could appear geminated in intervocalic position. This interpretation is not hugely inconsistent with the phonetic contrast—[r] and [ɾ] are both voiced apico-alveolar rhotics, with the trill of longer duration than the tap. The basic advantage of this interpretation as single vs. geminate, though, is that the distribution of contrast (in intervocalic position) and neutralization (elsewhere) falls out automatically. The contrastive trill occurs just where other geminate, moraic sonorants are found in contrast with their non-moraic counterparts (15a). No other context permits a moraic–non-moraic contrast, given that Catalan allows (with only the most narrowly defined exceptions) only one (moraic) consonant in a rhyme position (see Chapter 7).

(15) (a)		m	n	l	ʎ	'r' (/R/)
	#_	*molt* 'many'	*nou* 'nine'	*lema* 'slogan'	*llet* 'milk'	*roure* 'oak'
	.C_	*	*	*pluja* 'rain'	*	*prou* 'enough'
	C._	*esma* 'instinct'	*digne* 'worthy'	*orla* 'fringe'	*	*honra* 'dignity'
	V._V	*cama* 'leg'	*nena* 'girl'	*vila* 'town'	*illa* 'island'	*cara* 'face'
	V_._V	*gemma* 'gem'	*tarannà* 'character'	*til·la* 'lime tea'	*ratlla* 'line'	*carrer* 'street'
	V_C	*sembla* 'seems'	*nansa* 'handle'	*belga* 'Belgian'	*vullga* 'want.3SG. PRS.SUBJ' (Valencian)	*tort* 'twisted'
	V_#	*llum* 'light'	*segon* 'second'	*arrel* 'root'	*all* 'garlic'	*futur* 'future'

(b)	*summe*	*bipenne*	*afil·le*	*rotlle* 'roll'	*corre*
	'supreme'	'two-winged'	'leafless'		'run.2SG.IMP'

From this perspective, it remains somewhat surprising that in several contexts single non-contrasting /R/ is realized as a trill, not as a tap. However, from the analysis of intervocalic [r] as a geminate sonorant /RR/, two other facts about the distribution fall out.

Stem-final /RR/ is not syllabifiable as a coda (since it consists of a sequence of consonants of which the second is not less sonorous than the first); just as other stem-final geminates are not (15b). When no other morpheme follows, epenthesis is required, as in (15b).[7] If one takes /r/ ([r]) as a non-geminate, the fact that it cannot make a single-consonant coda appears mysterious. No other consonant is excluded from word-final position, except for /ʒ/, which is strengthened to [d͡ʒ] in word-final position.

In fact, the parallels between the distribution of [r] – [ɾ] on the one hand and [d͡ʒ] – [ʒ] on the other—see §2.1.3.3 above—have been remarked on by Mascaró (1987b), and taken up by Lloret (1992). In the case of [ʒ] and [d͡ʒ] too, the contrast is available only in intervocalic position, where there are some reasons to think that [d͡ʒ] is a geminate affricate. Non-contrasting /ʒ/ is, like /ɾ/, subject to fortition in word-final position, though categorically rather than variably; and it is subject also to fortition in word-initial and post-sonorant position, though variably, rather than categorically. The fact that in both cases the 'stronger' of the pair appears, to some extent, in neutralizing environments may reflect the fact that in some sense these particular 'strong' consonants, [d͡ʒ] and [r], are unmarked (relative to their 'weaker' counterparts [ʒ] and [ɾ]) in certain environments, namely, 'edge' environments: word-initially, word-finally, and in post-consonantal onset. Bonet & Mascaró (1997) object that the number of lexemes showing epenthesis of the unmarked vowel after '/ɾɾ/' is small. They claim only two items: *esquerre* 'left', and *corre* 'run' as in (15b) above—though this latter item shows consistent epenthesis in several inflectional contexts, not only finally (2SG.IMP, 3SG.PRS.IND) but also /+z/ *corres* (2SG.PRS.IND) and /+g/ in several varieties: *córrec* (1SG.PRS.IND), *córrega* (1/3SG.PRS.SUBJ), etc. One can add *parterre* 'flowerbed', *escirre* 'scirrhus', and *torre* 'tower' whose final [-e] in western Catalan cannot be identified with the feminine gender affix (as the final [-ə] in eastern Catalan might be). But the real point is there are no items in which '/RR/' is allowed to stand in final position.[8] Most masculine nominal roots ending in '/RR/' in fact take the

[7] In the case of /m/, /n/, and /l/, at least, the examples are few and somewhat recherché; the point is nevertheless that, insofar as stems end in geminate sonorants, epenthesis is necessary: *[summ], *[bipɛnn], *[əfill], *[rɔʎʎ] are impossible.

[8] Colomina (1996: 217) draws attention to a few possible cases: *esquer*, southern Valencian dialectal form of *esquerre*, *cigar* 'cigar', an artificial dictionary form for *cigarro*, whose derivatives have intervocalic [r], and *tor,* a form of *torre* 'tower' found only in place names such as *La Tor de Querol*.

masculine /+o/ suffix. There are a good score of these, including *burro* 'donkey', *carro* 'cart', *ferro* 'iron', *esguerro* 'bungling', *gerro* 'jug', *llobarro* 'sea-perch', and *porro* 'leek'. The final [ř] ∼ [ɾ] that occurs in words like *futur* 'future' (15b) always corresponds to intervocalic [ɾ] in derivatives, such as *futura* [fuˈtuɾə] 'future.F'.

The other fact that falls out from taking intervocalic [r] as realizing a geminate is one relating to stress constraints. Unlike most consonants, [r] is not found after the post-tonic syllable of a proparoxytone. *Còlera* [ˈkɔləɾə] 'cholera' is well-formed, but *[ˈkɔləɾə] would not be. The other non-moraic (onset) consonants excluded from this position are all palatals: /ʃ/, /ʒ/, (/tʃ/, /dʒ/), /ɲ/, /ʎ/, unlike [r]. However, if one takes intervocalic [r] to be a geminate, that is, mora-bearing, its exclusion from the post-tonic syllable of a proparoxytone follows automatically. A heavy pre-final syllable must bear lexical stress, if the final syllable itself does not (§9.5). Note that clusters which are onsets do not attract stress, so *càtedra* [ˈka.tə.ðɾə] 'university chair' is well-formed, as is *múltiple* [ˈmul.ti.plə] 'multiple'. Neither the final vowel epenthesis case for /RR/ nor the evidence of stress patterns is overwhelming; their significance is that independently they point in the same direction.

The account offered by Bonet & Mascaró (1997) primarily attempts to explain the distribution of the rhotic types in the non-contrastive contexts, developing approaches to syllabification and the sonority hierarchy derived from Clements (1990). The central ideas are that [r] is significantly less sonorous than [ɾ], and that [r] is unmarked, relative to [ɾ]. However, the claim that intervocalic [r] is the unmarked member of the trill/tap contrast in Catalan is in conflict with the evidence of sonorant dissimilation adduced by Lloret (1997). When a rhotic is the outcome of dissimilation, it is the tap, not the trill, which is found. Historical cases are *nonanta* > *noranta* 'ninety', *conomina* > *coromina* 'open field', LILIUM > *lliri 'lily'*; current non-standard pronunciations are *bona nit* > [ˈbɔɾəˈnit] 'good night', *juliol* > [ʒuˈɾjɔl] 'July'. Dissimilation shifting a rhotic to another sonorant also affects [ɾ], not [r], as Lloret's examples show. Bonet & Mascaró's account depends on their taking [r] as equivalent to an obstruent and [ɾ] as equivalent to a glide for sonority sequencing purposes, though these sonority equivalences are not independently argued for. On this basis [r] is preferred as a sole onset consonant, and [ɾ] is preferred in second position in an onset cluster, since it makes an appropriate gradient between an obstruent and a vowel. On the same sonority basis, [ɾ] is preferred in codas. As for the 'short trill' [ř] in codas, they merely say (p. 113) 'there is a postlexical process of tensing which applies to these coda [ɾ]s'. The contrastive intervocalic tap [ɾ], which is in onset position, for example, in *mira* [ˈmi.ɾə] 'look-at.3SG.PRS.IND', deviates from this putative unmarked distribution. For Bonet & Mascaró it is distinguished by being marked as [+f(lap)] in the underlying representation. That is to say, all non-contrasting rhotics, and intervocalic trills, are represented underlyingly as a rhotic /R/,

unspecified for the feature [flap], while the contrasting tap is represented as /$R_{[+flap]}$/. As mentioned above, Bonet & Mascaró discount the argument from epenthesis for intervocalic [r] being /RR/ on the grounds that it is supported by few cases. They discount the evidence from stressability, mentioned in the previous paragraph, on the grounds that [r] shares this property with the palatals, for which a geminate interpretation is implausible.

In section §4.2 of his thesis, Serra (1996b, expanding somewhat the account in Serra 1996a), makes three proposals for Catalan rhotics. The tap /ɾ/ is taken to be underlying. It is subject to an Optimality Theoretic constraint GEMINATE [ɾ] IN ONSETS, for which Serra does not offer independent motivation. (It seems typologically rather implausible, as formulated.) The constraint GEMINATE [ɾ] IN ONSETS is ranked below *COMPLEX, which penalizes margins with more than two segments. For *prec* 'request', this constraint ranking correctly prefers [pɾek] to *[prek]. The contrast between trill and tap in intervocalic position is made to depend on representing the intervocalic tap as in coda position in the input. That is to say, Serra's approach requires underlying syllabification of some forms. Thus *cera* 'wax' is underlyingly /ˈsɛɾ.ə/, as opposed to *serra* 'saw' /ˈsɛɾə/. In *serra* but not in *cera* the underlying /ɾ/ will surface as a trill due to GEMINATE [ɾ] IN ONSETS. (Serra's constraint ranking (251), FAITH » *COMPLEX » GEMINATE [ɾ] IN ONSETS, must be wrong since with underlying /ˈsɛɾə/ for *serra* it will prefer [ˈsɛɾə] to [ˈsɛrə]; no 'geminate' [r] will surface here or anywhere. Serra does not discuss how his syllabified representation /ˈsɛɾ.ə/ is supposed to interact with syllable structure constraints such as ONSET and *CODA.) The variable fortition of /ɾ/ in codas is informally related by Serra to processes which strengthen the right edge of various prosodic domains, but he offers no precise formulation.

Rhotic contrasts in a number of languages are addressed by Bradley (2001), though Catalan is not treated in any detail. However, he gives a good deal of attention to the typology of contrasting trill and tap, with phonetically based arguments for their distribution when this is not free. He uses Spanish as one of his example languages, in which language the distribution of trill and tap is sufficiently similar to Catalan to make several aspects of his account transferable. Several of the languages Bradley investigates allow contrast between trill and tap in a wider range of environments than Spanish does (or Catalan). However, the typology indicates three phenomena that are excluded. If there is neutralization in word-initial position, the trill, not the tap is found; thus #ɾ ⊃ #r. Thus, in initial position the trill is unmarked. Secondly, there is no contrast between trill and tap in homorganic clusters, although, for example, [rn] and [ɾn] may be in free variation (as they are in Catalan). Thirdly, following on from the previous observation, after homorganic sonorants only the trill is found: thus *[nɾ], *[lɾ]. Bradley's discussion of the data implies that he takes 'homorganic' to mean 'alveolar' in relation to the alveolar taps and trills that are the focus of his

investigation. But his more formal account of the *FAST/SAME SITE constraint which deals with the phenomenon does not distinguish between alveolars and dentals (only the feature [cor] is mentioned[9]). Thus his account remains vague concerning what is observed or expected with dental+rhotic or rhotic+dental sequences. Descriptions of Catalan generally distinguish between dental+rhotic clusters which are onsets -.tr- (*altre* 'other'), -.dɾ- (*cendra* 'ash'), -.ðɾ- (*pedra* 'stone'), always with taps, and alveolar+rhotic clusters which are coda+onset -n.r- (*honra* 'dignity'), -l.r- (*folro* 'lining'), -z.r- (*Israel* 'Israel'), always with trills. But such a prosodic distinction is not available to Bradley, who rejects a 'licensing-by-prosody' approach in favour of a 'licensing-by-cue' one, after Steriade (1997). A strict dichotomy between 'licensing by cue' and 'licensing by prosody', such that one is always to be preferred on theoretical grounds, does not seem well-founded, however. There is no reason to suppose speakers do not have access to knowledge both of articulatory and auditory phonetic cues on the one hand and of levels of prosodic structure, including syllable structure, on the other.

Bradley's account is based on an articulatory description of taps and trills formulated by him as follows (2001: 8–9):

> Whereas the alveolar tap is characterized by an extra-short closure duration of approximately 20 ms for Castilian Spanish, the alveolar trill has a longer, sustainable duration of 85 ms with three occlusions on average.[10]... Both the approach and release phases are crucial for successful articulation of the ballistic tap. On the other hand, the trill requires a tensed, controlled, and precise gesture in order to initiate passive vibration of the articulator by virtue of the Bernoulli effect. The trill is not simply a sequence of taps; the two rhotics involve completely different production mechanisms... With regard to perception,... cross-linguistically, taps tend to prefer intervocalic position and avoid word edges in order to maintain voicing and enhance perceptibility... The tap's brief closure duration is cued by the rapid transitions between its alveolar contact period and the adjacent vowels. On the other hand, the trill can be said to possess internal durational cues, since its duration may be sustained. This contrasts with the alveolar tap, which requires some degree of surrounding sonority to ensure perceptibility of its extra-short closure duration.

Because perceptibility of the tap is especially favoured intervocalically, Bradley argues that intervocalic position is the unmarked site for trill/tap contrast, even

[9] Bradley's *FAST/SAME SITE is formulated thus (11): 'Avoid faster-than-usual articulatory transitions in clusters involving the same articulator at the same constriction site', and is illustrated with a representation showing [nr] more harmonic than [nɾ]. As formulated, though, it would also always prefer [rn] over [ɾn], which is not what Bradley intends, apparently.

[10] Recasens (1991a: 267–8) gives the following figures for Catalan: [ɾ] 20–30 ms contact; [r] 2 to 3 contacts of 15–20 ms each, with gaps between the contacts of about 30 ms. Thus the minimum length for a 2-contact [r] is 60 ms, and for a 3-contact [r], 105 ms.

though for most consonants the perceptual salience of word-initial position favours that site as the unmarked position for contrasts.

Bradley proposes two kinds of articulatory markedness constraints relevant to apical constrictions: *FAST and *HOLD[11]. Context-free *FAST (16) militates against taps. (It is essentially equivalent to *ɾ.) Context-free *HOLD (17) militates against trills (and is equivalent to *r).

(16) *FAST: Avoid faster-than-usual articulatory transitions.

(17) *HOLD: Avoid a longer constriction.

These constraints reflect a typology of rhotics such that one can denote a trill [r] as 'rhotic[hold]' and a tap [ɾ] as 'rhotic[fast]'. (A rhotic approximant that is neither a trill or a tap, namely [ɹ], we may call 'rhotic[lax]'.) The unmarked nature of taps in intervocalic position can be formulated as a contextualized *HOLD constraint as in (18).

(18) *HOLD/V_V: Avoid a longer constriction in intervocalic position.

Similarly the unmarked nature of taps in complex onsets can be formulated as in (19).

(19) *HOLD/[$_\sigma$C_ : Avoid a longer constriction in a complex onset.

By the constraint *HOLD/[$_\sigma$C_ , onset clusters [.tɾ], [.βɾ], [.fɾ], and so on, are evaluated as more harmonic than [.tr], [.βr], [.fr], respectively. The typological markedness of initial taps can be expressed by the constraint *FAST/initial (20), after Bradley (2001: 11). From this point on, the analysis I offer deviates markedly from the details of Bradley's approach.

(20) *FAST/initial: Avoid faster-than-usual articulatory transitions in initial position.

I propose to pursue here the traditional GP analysis of the intervocalic trill/tap contrast (*serra* 'saw' versus *cera* 'wax') by representing the contrast as one of a geminate vs. a non-geminate rhotic (whose other characteristics need not be specified, as they will be governed by OT output constraints). In line with this, I adopt a phonetic transcription of the intervocalic trill as [r.r]; such a transcription emphasizes the point that the contrastive trill is both long (with respect to [ɾ]) and moraic. An alternative approach might be to represent the contrastive trill as 'rhotic[hold]' and make it subject to a faithfulness constraint (viz. IDENTRhotic[hold]) outranking *HOLD/V_V. But this would not help with

[11] For mnemonic reasons, I deviate from Bradley's usage and re-establish here the convention of using an asterisk for negative markedness constraints.

the epenthesis and stress facts mentioned above which point to contrastive [r] being moraic.

I can now turn to illustrating how the markedness constraints mentioned so far evaluate some alternative rhotic outputs. The examples are those given in (15) above. Where relevant, input rhotics will be represented /R/, standing for any trill, tap, or indeed lax approximant; by the theory of 'Richness of the Base' the choice between them is a matter of indifference. (Lexicon Optimization will prefer the surface rhotic as input, of course.) In (21) I consider alternatives for *roure* [row.ɾə] 'oak', which contains two rhotics. I assume a high-ranking constraint ruling out lax [ɹ] for most dialects.

(21)

roure	*HOLD/[$_\sigma$C_	*HOLD/V_V	*FAST/init
row.rə		*!	
☞ row.ɾə			
ɾow.rə		*!	*
ɾow.ɾə			*!

In (22) I consider alternative candidates for *prou* 'enough' and *cara* 'face'.

(22)

(a)

prou	*HOLD/[$_\sigma$C_	*HOLD/V_V	*FAST
prɔw	*!		
☞ pɾɔw			*

(b)

cara			
ka.rə		*!	
☞ ka.ɾə			*

For words like *tort* 'twisted', as mentioned in §2.1.6.1, one needs to allow for two pronunciations, one with a (short) trill and one with a tap in the coda, each of which is preferred in different varieties. I interpret these variants here as reflecting two alternative rankings of the basic *HOLD and *FAST constraints. When *FAST dominates, the trill is preferred; when *HOLD dominates, the tap is preferred. However, as seen above (13), before a non-strident fricative the tap is preferred in all varieties, e.g. *herba* [ˈeɾ.βə]. Thus, in the short-trill varieties where *FAST outranks *HOLD, *FAST must itself be outranked by a constraint disfavouring the trill before a non-strident continuant: *HOLD/__[−son, −strid, +cont]. Tableau (23) illustrates *tort* and *herba* in the two varieties, here simply labelled A and B.

(23)

A

tort	*HOLD/_ [–son, –strid, +cont]	*FAST	*HOLD
☞ tɔrt			*
tɔɾt		*!	
herba			
'er.βə	*!		*
☞ 'er.βə		*	

B

tort	*HOLD/_ [–son, –strid, +cont]	*HOLD	*FAST
tɔrt		*!	
☞ tɔɾt			*
herba			
er.βə	*!	*	
☞ er.βə			*

I can now turn to the case of the intervocalic trill, illustrated in (24) with *carrer* 'street', where the first constraint, MAXμ, penalizes deletion of moras.

Tableau (24) illustrates how the intervocalic trill reflects the operation of the same constraints as govern the distribution of trill and tap in non-contrasting environments. On the model of tableau (24), in the *HOLD » *FAST dialects, the version [kər.ˈre], represented with a coda tap, would be preferred. It is perhaps more likely that a GEMINATE INTEGRITY constraint should exclude realizations like [ɾ.r] or [r.ɾ] in any case.[12]

(24)

carrer /kaRˈRe/	MAX	*HOLD/[σC_	*HOLD/V_V	*FAST	*HOLD
kə.ˈre	*!		*		*
kə.ˈre	*!			*	
kəɾ.ˈɾe				*!*	
kəɾ.ˈre				*!	*
☞ kər.ˈre					**
kər.ˈɾe				*!	*

[12] Padgett (2003) offers a phonetically and theoretically detailed account of Catalan rhotics and their distribution, in terms of Dispersion Theory, that builds on Bradley (2001). Padgett's paper came to my attention too late to be reflected in the exposition here. It offers many interesting proposals, though relying on two levels of constraint hierarchy, lexical and postlexical, an approach to the grammar of phonology that this book avoids.

Lastly, the case of *honra* [ˈon.rə] 'dignity' (15a) requires attention to a further complication. When a verb stem (conjugation II or III) that ends in an alveolar nasal or lateral is followed by inflectional /+R/—an inflectional morpheme found in the infinitive, and in the future and conditional—epenthesis of [d] occurs. Thus, from the root /enten/ 'understand', the 3SG.FUT /+R+ˈa/ gives *entendrà* [əntənˈdɾa]. From the root /bal/ 'be worth', the 3SG.FUT is *valdrà* [bəlˈdɾa].[13] As *honra* [ˈon.rə] shows, epenthesis of [d] does not occur within morphemes. I return to this matter in §8.2.

The question of 'final *-r* deletion', which is subject to dialect difference and is lexically conditioned, will be treated in §10.2.2. For example, the final consonant of *futur* 'future' is pronounced in all varieties, as a tap or a short trill as the case may be, while *carrer* is pronounced [karˈreɾ] in central Valencian, but without a final consonant in other varieties. In such words in the other varieties the root-final /-ɾ/ surfaces only prevocalically, in derivatives like *carreró* 'alley'.

2.1.7 *Laterals*

2.1.7.1 *Alveolar lateral*

Catalan /l/ is notably velarized (to a higher degree after back vowels or in coda contexts, though not exclusively there: Wheeler 1979: 306–8). This is one of the most noticeable characteristics of Catalan speech which differentiates it from Spanish (or French) even for listeners unfamiliar with the languages. Except where the velarized quality of /l/ is in focus (e.g. Chapter 6), the unmodified symbol [l] is used in this book.

2.1.7.2 *Alveolo-palatal lateral*

The alveolo-palatal lateral /ʎ/ participates in two patterns of sociolinguistic variation (Pons 1992: 12–24). These patterns are similar in their phonetic effects—in both patterns [ʎ] is in variation with [j]—but they differ in their history and sociolinguistic distribution. The first, traditional, pattern concerns the lexical incidence of [ʎ] and [j]. There is one set of words in which [ʎ] is the norm in all regional varieties. This set consists of words such as those in (25a) in which [ʎ] descends from Latin initial L- or intervocalic -LL-.

(25)	(a)	invariant /ʎ/	(b)	/ʎ/ ∼ /j/
		lluna [ˈʎunə] 'moon'		*all* 'garlic'
		llegir [ʎəˈʒi] 'read'		*fulla* 'leaf'
		cavall [kəˈβaʎ] 'horse'		*tallar* 'cut$_V$'
		ella [ˈeʎə] 'she'		*orella* 'ear'
				ull 'eye'

[13] Some dialects in the north lack d-epenthesis, and have forms such as *entenrà* [əntənˈra], *valrà* [bəlˈra], in line with the morpheme-internal pattern of rhotic distribution seen in *honra*.

The other set consists of words which had -LI̯- or -C'L-in Latin, such as those in (25b). The words in this set have [ʎ] in the majority of regions (and this pronunciation is represented in the standard orthography). In two areas, though, [j] is found in the words of the (25b) set, namely, in the Balearic Islands and in much of north-eastern Catalonia, east of the river Llobregat, north and east of Barcelona. The pronunciation involving [j] in words of this set is known in Catalan as *ieisme*, and has been in existence for many hundreds of years. Within the Balearic Islands *ieisme* is firmly established and is regarded as properly characteristic of the regional speech variety: such words as those in (25b) are there regularly pronounced [ˈaj], [ˈfujə] [təˈja], [oˈɾəjə], [ˈuj].[14] On the mainland, however, *ieisme* is in recession. It is stigmatized as a marker of older, rural, less educated speakers. Pons's investigation (1992) in 1983 in Barberà del Vallès, a town some 20 km from Barcelona where *ieisme* is part of the traditional dialect, revealed a marked decline in the number of [j] realizations in the relevant word class among speakers born after 1930, except in the handful of words where [j] is current in Barcelona speech. (These are words such as *cella* [ˈsɛjə] 'eyebrow', *ullal* [uˈjal] 'eye tooth', *llentilla* [ʎənˈtiə] 'lentil' (with [ij]→ [i] see §3.2.5.2).) From the phonological point of view, traditional *ieisme* is a matter of dialect difference in the lexical distribution of /ʎ/ and /j/. On the mainland, *ieista* dialects are succumbing to the prestige of the dialect of the capital in this respect, as in others.

The second pattern is a much more recent tendency to replace all instances of [ʎ] by [j]. This has been remarked on since the 1970s as characteristic of young speakers in several areas (Recasens 1991b: 324), but has not, as far as I know, been the subject of systematic investigation. Pons (1992) observed that it was not part of the speech of any of her informants in Barberà del Vallès in 1983, and her informants did not then recognize her taped example of it as normal Catalan. It was also absent in Pi-Mallarach's (1997) investigation of students in Vic, born in the mid-1970s. Those who have remarked on this development have taken it to reflect Spanish influence. In 'non-standard' Spanish shift from /ʎ/ to /j/, involving phonemic merger, is widespread. Of course, a shift from a marked sound type, such as [ʎ], to a less marked one, such as [j], is not unexpected. Within Europe such a change has taken place in Romanian, Hungarian, French, and Occitan (and must also have taken place in Catalan a millennium or so ago to give rise to traditional *ieisme*). The 'actuation' question invites investigation, though. Why did this development take root in Catalan in the last quarter of the twentieth century in particular?

[14] Or in intervocalic position there may be lowering of /j/ to [e̯], or loss of /j/, which is characteristic of Minorca.

2.1.8 *Place in coronals*

Not unusually, the dental coronals (/t/, /d/) and the alveolar coronals (/n/, /l/, /R/, /s/, /z/) of Catalan are subject to place assimilation to each other and to other adjacent coronals, alveolo-palatal and dorso-palatal. The details are reviewed in Chapter 6.

2.1.9 *Contrasting geminate consonants*

In §2.1.3.3 geminate [dd͡ʒ] contrasting with [d͡ʒ] ∼ [ʒ] was considered. Contrasting geminate obstruents are rather rare in Catalan, though there are some examples (26) of /dd/ and /bb/, not all of which are plausibly bimorphemic.

(26) (a) *addenda* [ədˈdendə] 'addenda' *adduir* [ədduˈi] 'adduce'
addicte [ədˈdiktə] 'addicted'
(b) *obvi* [ˈɔbbi] 'obvious' *obvers* [ubˈbɛrs] 'obverse'

Geminate sonorants are rather more common, though */ɲɲ/ is not attested. (The interpretation of intervocalic [r] as /RR/ is controversial, though defended here; see §2.1.6.1.) Examples of geminate sonorants are given in (27).

(27) (a) *gemma* 'gem' ≠ *gema* 'crystallized yolk' *setmana* [səmˈmanə] 'week'
summa 'summary' ≠ *suma* 'sum' *ametlla* [əmˈmɛʎʎə] (sic) 'almond'
(b) *cotna* [ˈkonnə] 'rind' *decenni* 'decade'
tarannà 'character' *nannar* (baby-talk term for 'walk')
(c) *til·la* [ˈtillə] 'lime tea' *taral·lejar* [tərəlləˈʒa] 'hum.INF'
atles [ˈallɛs] 'atlas' *xarel·lo* [ʃəˈrɛllu] (grape variety)
atleta [əlˈlɛtə] 'athlete'
(d) *motlle* [ˈmɔʎʎə] 'mould' *espatlla* [əsˈpaʎʎə] 'shoulder'
vetllar [bəʎˈʎa] 'stay awake' *bitlla* [ˈbiʎʎə] 'skittle'

Geminate /ʎʎ/ as illustrated in (27d) is not found in Valencian or Balearic, or in south-west Catalonia; these varieties have /ll/ instead in the corresponding words, with spelling to match: *motle*, *vetlar*, etc. Some varieties, such as Valencian *apitxat* avoid geminate sonorants altogether (promoting a *GEMINATE constraint). Normatively, the orthographic sequences <*nn*> and <*l·l*> are always supposed to correspond to geminate pronunciation, but such pronunciations rarely occur spontaneously in the Greco-Latin vocabulary elements. So *mil·lenni* 'millennium' is more often [miˈlɛni] than [milˈlɛnni]. Orthographic <*tm*>, <*tn*>, <*tl*>, <*tll*> quite neatly reflect the fact that the first part of the geminate in each case could be seen as a consonant unspecified for place, voice, and laterality or nasality. Or, to put it another way, [mm], for example, is not distinct

from /tm/, /dm/, /pm/, /bm/, or /nm/; however, /gm/, /lm/, /zm/ remain distinct; see §6.3.

The geminate realization of /b/ and /g/ before /l/ in certain contexts is discussed in §8.5. There is a marginal contrast between [gg] and [ɣ] in *segle* [ˈsegglə] 'century', *iglésia* [igˈglezjə] 'church' (non-standard for *església*) vs. *sigla* [ˈsiɣlə] 'acronym, initialism', *ègloga* [ˈɛɣluɣə] 'eclogue', but these are usually treated as isolated exceptions.

2.2 VOWEL INVENTORY

Among the most striking differences between Catalan dialects are differences in the number and incidence of vowel phonemes in stressed syllables, and in the nature and extent of reduction in vowel contrasts in unstressed syllables (Julià 2002; Jiménez 2002). In this section aspects of vowel contrasts in stressed syllables are considered. Vowel reduction is treated in §2.3.

2.2.1 *Stressed vowel contrasts*

Table 2.2 presents the maximal inventory of vowel contrasts found in stressed syllables. This eight-member inventory is characteristic only of Balearic Catalan.

In (28) is a quasi-minimal set illustrating the contrasts in the Balearic variety, in the environment before final /k/.

(28) /i/ *ric* 'rich' /u/ *suc* 'juice'
/e/ *séc* 'fold' /ə/ *sec* 'dry' /o/ *moc* 'snot'
/ɛ/ *sec* 'sit.1SG.PRS.IND' /a/ *sac* 'bag' /ɔ/ *soc* 'clog'

The stressed vowel /ə/ is absent from continental varieties of Catalan. As a broad generalization, we can say that words that have /ə/ in Balearic have /ɛ/ in eastern continental Catalan, and /e/ in western Catalan. Thus *sec* 'dry' is pronounced /sɛk/ in eastern continental Catalan (homophonous with *sec* 'sit.1SG.PRS.IND'),

TABLE 2.2 Inventory of stressed vowel types

	Front	Central	Back rounded
Close	i		u
Half-close	e	ə	o
Half-open	ɛ		ɔ
Open		a	

but /sek/ in western Catalan (homophonous with *séc* 'fold', also with *cec* 'blind', and, indeed with *sec* 'sit.1SG.PRS.IND', which also happens to have /e/ in those varieties). North Catalan, though regarded as part of the eastern dialect block, has merged /e/ and /ɛ/ as a mid front vowel usually represented as half-close /e/ (so *séc* 'fold', *cec* 'blind', *sec* 'sit.1SG.PRS.IND', and *sec* 'dry' are homophonous in north Catalan too). North Catalan has also merged /o/ and /u/ as /u/, so *moc* 'snot' [muk] and *suc* 'juice' [suk] contain the same vowel phoneme; consequently the remaining non-high rounded vowel in north Catalan, with lexical incidence corresponding to /ɔ/ in other varieties, is not especially open: *soc* 'clog' can be represented as [sok] /sok/ there. North Catalan, then, has a familiar five-vowel inventory, /i, e, a, o, u/, in stressed syllables.

If we leave aside Balearic Catalan with an eight-vowel inventory, and north Catalan with a five-vowel inventory, the remaining dialects have a familiar inventory with seven peripheral vowels: /i, e, ɛ, a, ɔ, o, u/. But central Catalan (i.e. eastern continental Catalan excluding north Catalan) and western Catalan differ markedly in the lexical incidence of the mid front vowel phonemes /e/ and /ɛ/. As mentioned above, words with /ɛ/ in central Catalan (corresponding to /ə/ in Balearic) generally have /e/ in western Catalan. (Words with /e/ in Balearic almost always have /e/ in central and western Catalan as well.) A few examples are given in (29).

(29)

	Western	Central	Balearic
plena 'full (F.SG)'	ˈplena	ˈplɛnə	ˈplənə
deu 'owe.3SG.PRS.IND'	ˈdew	ˈdɛw	ˈdəw
set 'thirst'	ˈset	ˈsɛt	ˈsət
ven 'sell.3SG.PRS.IND'	ˈven	ˈbɛn	ˈvən

One consequence of this pattern of incidence is that whereas in central Catalan the lexical frequency of /e/ and /ɛ/ is fairly evenly balanced,[15] in western Catalan /e/ predominates over /ɛ/. For example, in Valencian /ɛ/ is very rare both in final position (perhaps only in *qüiè* 'group of four' and *Novetlè* (toponym)), and before final /t/ (perhaps only in *set* /sɛt/ 'seven', which makes a minimal pair with *set* /set/ 'thirst': Renat i Ferrís 1943).

2.2.1.1 *Gaps in the distribution of mid front vowels in central Catalan*

Whereas restrictions have not been noted on the distribution of /ɛ/ in central Catalan, there are some curious gaps in the distribution of /e/, depending especially on the consonant or consonant sequence following.

[15] In a text corpus of 10,254 words Rafel (1981) observed stressed vowels in the following proportions (central Catalan pronunciation): /i/ 15.73%, /e/ 14.52%, /ɛ/ 10.08%, /a/ 29.99%, /ɔ/ 8.86%, /o/ 10.93%, /u/ 9.90% (figures calculated directly from Quadre IV in Rafel 1981: 480–84). Text frequency is likely to differ somewhat from lexical frequency, but I know of no counts directly of the latter.

(30) Gaps and contrasts in the distribution of /e/ and /ɛ/

		/e/	/ɛ/
(a)	before /l/	*	*cel* /sɛl/ 'sky'
(b)	before R[+coronal]	*	*obert* /obɛRt/ 'open'
(c)	before [r]	*	*terra* /tɛRRa/ 'earth'
(d)	before R [−coronal]	*herba* /ˈeRba/ 'grass'	*verba* /ˈbɛRba/ 'jest'
		terme /ˈteRme/'boundary'	*isoterm* /izoˈtɛRm/ 'isotherm'
		cercle /ˈseRkle/'circle'	*tubercle* /tuˈbɛRkle/ 'tuber'
		Berga /ˈbeRga/(toponym)	*verga* /ˈbɛRga/ 'stick'

Firstly, /e/ does not precede /l/ (30a).[16] The historical sources of /e/ in other contexts derive only /ɛ/ before /l/. In a comparable way, the sources of /e/ in other contexts derive only /ɛ/ before a rhotic ([r] or [ɾ]; see §2.1.6) followed by a [+coronal] consonant (30b), or before /R/ (30c). (If one interprets /r/ as geminate /RR/ as in §2.1.6, these two contexts become one and the same.) Note that /e/ and /ɛ/ do contrast before /R/ followed by a [−coronal] consonant, as in (30d). In fact, the absence of /e/ before /l/, before /R/ followed by a [+coronal] consonant, and before [r] is virtually universal in Catalan, though in Balearic /ə/ occurs in some words (< Latin /i/ or /eː/) corresponding to /ɛ/ in other varieties: *pèl* Bal. /ˈpəl/, central and W. Catalan /ˈpɛl/, 'hair'; *verd* Bal. /ˈvəRd/, central and W. Catalan /ˈbɛRd/ 'green'; *cerra* Bal. /ˈsəRRa/, central and W. Catalan /ˈsɛRRa/ 'bristle'. The articulatory reasons for this skewed distribution are not entirely apparent. Catalan /l/ is characteristically velarized to a greater or lesser degree, so anticipation of raising of the back of the tongue, for /l/, could lead to lowering of the front of the tongue in a preceding mid vowel. But one might expect a similar effect on the sequence /il/, where the target tongue positions for the vowel and the consonant are in conflict. It is true that the sequence /il/ is extremely rare before a consonant in Catalan,[17] but /il#/ and /ilV/ are quite well attested (e.g. *abril* 'April', *concili* 'council', *vila* 'town', *quilo* 'kilo'). While recognizing the 'difficulty' of the sequence consisting of a close or half-close front vowel followed by velarized /l/, it may be that no alternative to [il] is regarded as more acceptable (more 'harmonic' in the OT sense). In the case of /el/, however, [ɛl] may have been allowed as an acceptable deviation, given the greater insecurity of the contrast between /e/ and /ɛ/ in general (see §2.2.1.3 below).

In the case of front vowels preceding rhotics, anticipation of the lowered tongue front required for the apico-alveolar tap or trill could well disfavour

[16] Some sources give *feltre* 'felt' with /e/. Badia i Margarit (1970) mentions that /ɛ/ is also found in this word. (*Peltre* 'pewter' has /ɛ/.) The name of the letter *L* is also usually represented as [ˈelə], but the current form of the letter names displays significant Spanish influence (Rossich 2000; see also /e/ in *F* [ˈefə], and *N* [ˈenə]; only *R* [ˈɛra] overcomes the tendency to maintain /e/ as the stressed vowel in the recent names of the consonant letters derived from their sounds).

[17] At least these exist: *filtre* 'filter', *quilma* 'large sack', *silf* 'sylph'.

close front vowels. In this case, unlike that of /l/, the 'problem' is more acute with [r] or /RC/ than it is with intervocalic [ɾ], since intervocalic position allows a fast targeting and withdrawal of the tongue tip whatever the position required for the adjacent vowels; for the trill or for a rhotic before another consonant involving the front of the tongue, a preceding close or half-close front vowel allows rather little space for organizing the controlled vibration of the tongue tip. Again, in principle, the same problem should of course beset /i/ before apico-alveolar rhotics. And, in fact, in Catalan /i/ is very rare before intervocalic [r] or a rhotic followed by a [+coronal] consonant.[18] From an OT perspective we would want to say that there are constraints, for articulatory reasons, against [+high] front vowels[19] before [ɫ], and before /R.[+coron]/ (*[+hi, −bk] ɫ, *[+hi, −bk] R.[+coron]). These are outranked by the faithfulness constraint requiring maintenance of the place features of /i/ ($\textsc{Ident}_{PA}$i), but not by other $\textsc{Ident}_{PA}$ constraints. Half-open /ɛ/ then surfaces due to minimal constraint violation.

In central Catalan half-close front /e/ is also avoided in favour of /ɛ/ in the following contexts:

before the velar nasal [ŋ], except, for some speakers only, in the word *llengua* 'tongue, language'
before the palatal nasal /ɲ/,[20] except in the words *lleny* 'boat, lignum' and *llenya* 'firewood'
before /f/
before /sk/, except, for most speakers, in the word *llesca* 'slice'

Avoiding /e/ before a velar or palatal nasal makes some sense as an auditory phenomenon. In the somewhat obscured formant structure of a phonetically nasalized vowel, raising the front part of the tongue for [e] may be interpreted as coarticulation and perceptually 'corrected'[21] to /ɛ/, leading in due course to lexicalization of /ɛ/ exclusively in this environment. I offer no phonetic account for the absence of */ef/ and */esk/. One thing that does become evident from the preceding discussion and examples is the raising effect of a preceding palatal lateral /ʎ/; the effect is clearly coarticulatory (and perseverative). Its presence is, for some speakers (Badia i Margarit 1970: §§9, 22), sufficient to outdo the lowering effect of the velar nasal in *llengua* 'tongue' (though not in *llenca*

[18] I have identified only *esmirle* 'merlin' (for which the variant form *esmerla* [əzˈmɛrlə] exists), *flirt* (Anglicism with spelling pronunciation), *mirra* 'myrrh', and *sirte* 'sandbank'. Before non-coronals /iR/ occurs in slightly less abstruse words, e.g. *circ* 'circus', *firma* 'signature', *sirga* 'tow-rope'.

[19] For the sake of argument, taking [+high] to correspond to close and half-close vowels, [+low] to open and half open vowels, and [+mid] to half-close and half-open vowels.

[20] /e/ occurs before alveolo-palatal [nʲ] homorganic with a following consonant, in *diumenge* 'Sunday', in the stressed present tense forms of *menjar* 'eat', and, according to Castellanos & Castellanos (1979), in *revenja* 'revenge' and the present of the associated verb *revenjar* 'take revenge' (Bruguera 1990 gives [rəˈβɛɲʒə]). Recasens (1991b: 84–5) gives further examples of variant pronunciations illustrating the raising effect of adjacent consonants on /ɛ/.

[21] 'Hypocorrection' in the sense of Ohala (1989: 184–8).

[ˈʎɛŋkə] ‘rasher’; anticipation of the following labio-velar glide may be what tips the balance for *llengua*). The same would be true for all speakers in the case of *llenya* (and *lleny*, insofar as that word is known to speakers).[22] Badia i Margarit (1970: §§9, 22) reports that *llesca* ‘slice’ has /e/ generally in Catalan, though /ɛ/ can also be heard in Barcelona. Before /f/, however, even preceding /ʎ/ is insufficient to raise /ɛ/ to /e/: *llefre* [ˈʎɛfrə] ‘grimy’, according to standard sources; but see below §2.2.1.3 on the preference for [ˈʎefrə] among students in Vic. In several other contexts where the following environment would in principle allow either /e/ or /ɛ/, preceding /ʎ/ means /ɛ/ is excluded. These are listed in (31). The following labials and alveolo-palatals are articulatorily consistent with the raising effect of /ʎ/.

(31) *lleba* = *falleba* [(fə)ˈʎeβə] ‘window catch’: */-ʎɛba/, */-ʎɛbe/
lleva [ˈʎeβə] ‘levy’, and likewise present tense forms of *llevar* ‘take off’
llebre [ˈʎeβrə] ‘hare’: */-ʎɛbre/, */-ʎɛbra/
llepa [ˈʎepə] ‘toady’, and likewise present tense forms of *llepar* ‘lick’: */-ʎɛp-/
llémena [ˈʎeməna] ‘nit’: */-ʎɛma/, */-ʎɛme/
lleixa [ˈʎeʃə] ‘shelf’, *relleix* [rəˈʎeʃ] ‘ledge’: */-ʎɛʃ-/
lleig [ˈʎet͡ʃ] ‘ugly.M’, *lletja* [ˈʎed͡ʒə] ‘ugly.F’: */-ʎɛt͡ʃ/, */-ʎɛd͡ʒ-/

2.2.1.2 *Mascaró’s law*

Mascaró (1984a) drew attention to the morphophonemic process in Catalan whereby half-close vowels in stems become half-open when followed by another syllable involving an unstressed derivational suffix containing the close front vowel /i/. Here I refer to this process as ‘Mascaró’s law’ in honour of its discoverer. Some examples are given in (32).

(32)	*Homer* [uˈmeɾ] ‘Homer’	*homèric* [uˈmɛɾik] ‘Homeric’
	modest [muˈðest] ‘modest’	*modèstia* [muˈðɛstjə] ‘modesty’
	esfera [əsˈfeɾə] ‘sphere’	*esfèric* [əsˈfɛɾik] ‘spherical’
	cent [ˈsen] ‘hundred’	*cèntim* [ˈsɛntim] ‘cent’
	carbó [kəɾˈβo] ‘coal’	*carboni* [kəɾˈβɔni] ‘carbon’
		carbònic [kəɾˈβɔnik] ‘carbonic’
	divisor [diβiˈzo] ‘divider’	*divisori* [diβiˈzɔɾi] ‘divisive’
	falcó [fəlˈko] ‘hawk’	*falcònids* [fəlˈkɔnits] ‘Falconidae’

Since there are several suffixes with unstressed /i/, and many of them are very productive, the phenomenon in (32) could be illustrated with hundreds of examples. In Mascaró’s account (1984a) the process is specified to take place before /i/ in derivational but not inflectional suffixes. It is in fact only the subjunctive suffix /+i/ that does not cause lowering of half-close vowels in stems or preceding

[22] *Pellenya* ‘wire sieve’ has /ɛ/ according to Bruguera (1990); for *llenya* it may help that /ʎ/ is initial.

affixes. Thus, for example, the first and second conjugations show past subjunctive endings *-essis* (2SG), *-éssim* (1PL), *-éssiu* (2PL) and *-essin* (3PL), /+ˈes+i+s/ [-ˈesis], etc. In the context of present subjunctive forms, this fact gives rise to some surface contrasts between the vowel in stems before derivational /+i/ and the vowel in the same stem before inflectional (subjunctive) /+i/, as seen in (33).

(33) *imperi* [imˈpɛɾi] 'empire' *imperi* [imˈpeɾi] 'rule.1/3SG.PRS.SUBJ'
refrigeri [rəfɾiˈʒɛɾi] 'cold snack' *refrigeri* [rəfɾiˈʒeɾi] 'refrigerate.1/3 SG.PRS.SUBJ'
carboni [kəɾˈβɔni] 'carbon' *carboni* [kəɾˈβoni] 'carbonize.1/3 SG.PRS.SUBJ'

The exceptionality of the subjunctive affix /+i/ with regard to the effect of [i] on a preceding stressed vowel is surely to be attributed to paradigm uniformity constraints enforcing non-alternation of stressed vowels in verb stems where /e/ or /o/ is lexical and appears in the present indicative, and in the past subjunctive suffix /+es/, which is [-ˈes] when no other inflection follows, in the unmarked 1/3SG forms, e.g. *cantés* 'sing.1/3SG.PST.SUBJ'.

Pi-Mallarach (1997: 21–2) expands Mascaró's observations, and points out that half-open /ɛ, ɔ/, rather than half-close /e, o/, are found before unstressed /i/ consistently in 'neologisms' whatever their morphological structure, such as *confeti* [kunˈfɛti] 'confetti', *florilegi* [fluɾiˈlɛʒi] 'florilegium', *iterbi* [iˈtɛɾβi] 'ytterbium', *bronquis* [ˈbɾɔŋkis] 'bronchi', *hòlding* [ˈxɔldiŋ] 'holding company', *OVNI* [ˈɔbni] 'UFO'. But, in fact, there does not seem to be much reason to think that the phenomenon in question is constrained either by derivational structure or by 'neologism' (however the latter is defined). Here are some examples of 'old' words, all with half-open vowels, firstly with /ɛ/:

adverbi 'adverb'	*geni* 'genius'	*senderi* 'common sense'
bèstia 'animal'	*gremi* 'guild'	*sènia* 'water wheel'
col·legi 'school'	*imbècil* 'idiot'	*sèquia* 'irrigation channel'
dèria 'craze'	*medi* 'environment'	*sèrie* 'series'
emprèstit 'loan'	*misteri* 'mystery'	*seti* 'location'
espècia 'spice'	*neci* 'stupid'	*tebi* 'lukewarm'
estèril 'barren'	*nervi* 'nerve'	*tedi* 'tedium'
estri 'tool'	*nyèbit* 'urchin'	*tènia* 'tapeworm'.
evangeli 'gospel'	*premi* 'prize'	
èxit 'success'	*règim* 'diet'	

And a comparable set, all with /ɔ/:

Antoni 'Anthony'	*dosi* 'dose'	*nòmina* 'payroll'
bigoti 'moustache'	*enjòlit* 'in suspense'	*obi* 'manger'
bòfia 'blister'	*escòria* 'slag'	*odi* 'hatred'
bòlit 'tip-cat'	*falòrnia* 'tall story'	*oli* 'oil'
bòria 'plot'	*fer l'orni* 'act dumb'	*òliba* 'owl'

borni 'one-eyed'
bòvila 'kiln'
bròquil 'broccoli'
carquinyoli '(type of) biscuit'
clòtxina 'shell'
codi 'code'
còfia 'bonnet'
còlit 'wheatear'
colònia 'settlement'
còpia 'copy'
cossi 'washtub'
crònica 'chronicle'
dimoni 'devil'
divorci 'divorce'
fòtil 'utensil'
glòria 'glory'
gobi 'gudgeon'
hipòcrita 'hypocrite'
història 'story'
Jeroni 'Jerome'
Jordi 'George'
hòstia 'host'
lògica 'logic'
macedònia 'fruit salad'
manicomi 'mental hospital'
memòria 'memory'
moixoni 'shhh!'
negoci 'business'
noli 'charter'
ordi 'barley'
orri 'sheep milking parlour'
parròquia 'parish church'
propi 'own'
propòsit 'purpose'
pròxim 'next'
safanòria 'carrot'
sobri 'sober'
soci 'member'
somni 'dream'
sòpit 'drowsy'
sòtil 'ceiling'
tòria 'vine shoot'
tòtil 'toad'
vori 'ivory'

Words which are exceptions to this principle are very rare: *sépia* 'cuttlefish' has alternative forms *sèpia* and *sípia*; Bruguera (1990—central Catalan pronunciation) only gives the latter two forms. The /e/ in *església* 'church' may well be a Castilian feature; [ɛ] can certainly also be heard in this word. The popular form of this word is, in any case, *iglésia* [igˈglezjə] (Valencian [iˈɣlezja]), whose closeness to the form of Spanish *iglesia* is yet more evident. In Valencian the following are reported with /ˈe/ preceding /i/: *sépia* 'cuttlefish', *sénia* 'water wheel', *séquia* 'irrigation channel', *Dénia* (toponym), and *València* (all other words in *-ència* have /ɛ/). Renat i Ferrís (1943) also mentions *béstia* 'animal' and *série* 'series' for Valencian, but other sources give *bèstia* and *sèrie* for Valencian, as in other dialects. Notice that all the 'exceptions' involve the front vowel pair /e, ɛ/, and only in the context before /-Cia/ or /-Cie/. Setting aside these possible exceptions, and the specific case of the subjunctive inflection /+i/, the situation is thus that before a following syllable involving a high front vocoid the stressed vowel inventory is reduced to five contrasting members: /i, ɛ, a, ɔ, u/. By Mascaró's law, /e/ and /o/ in stems are lowered to /ɛ/ and /ɔ/ respectively in morphological derivations introducing this context. I suggest that a suitable explanation is similar to that proposed above for the neutralization of /e/ and /ɛ/ as /ɛ/ before the palatal nasal /ɲ/. That is to say, there is an element of dissimilation involved—a half-open vowel is less similar to a following /i/ than a half-close vowel would be—though more needs to be said so as to explain why close vowels /i/ and /u/ in stems are not subject to 'lowering' by dissimilation. In an assimilatory raising environment it is more difficult to maintain the auditory dispersion required to distinguish seven vowel qualities (eight in Balearic) perceptually. Anticipatory assimilative raising would involve a degree of overlap between the realizations of the seven vowels in this context and their realizations in 'freer' contexts. Historically, mishearing or misinterpretation probably did occur, especially between the pairs of mid vowels whose distinctiveness has long

been relatively insecure. As a consequence the phonemic distinction between half-open and half-close mid vowels may be abandoned in this context, favouring the lower of the pair, less likely to be misheard as a close vowel. The alternative forms for 'cuttlefish', *sèpia*, *sépia*, and *sípia*, may well be indicative of this very point; several other alternative forms with root /i/ as one of the alternatives are recorded in standard dictionaries, such as *bístia* = *béstia* = *bèstia* 'animal', *sínia* = *sénia* = *sènia* 'water wheel', *síquia* = *séquia* = *sèquia* 'irrigation channel', *nici* = *neci* 'stupid', *nirvi* = *nervi* 'nerve', *siti* = *seti* 'location', and Old Catalan *exempli* = *eximpli* 'example' (modern Catalan *exemple*).

Neutralization of mid vowel contrasts before /i/ in a following syllable may be just the most evident aspect, because virtually categorical, of a more general tendency, noted by Fabra (1908), by which half-open mid vowels /ɛ/ and /ɔ/ are strongly preferred in paroxytones and proparoxytones before rounded vowels /u/ and /o/, as well as before /i/, or before any vowel followed by a further consonant in the stem. The following list (34) is a fairly comprehensive one of words which deviate from this principle, having /e/ or /o/, as the case may be.[23]

(34)

cérvol 'deer'	*foto* 'photo'	*mosso* 'cabin boy'
coco 'coconut'	*furóncol* 'boil'	*moto* 'motor cycle'
déntol 'Dentex dentex'	*golfo* 'hinge'	*préssec* 'peach'
ergo 'therefore'	*góndola* 'gondola'	*préstec* 'loan'
escórpora 'scorpion fish'	*gotzo* 'corpulent'	*témpores* 'ember days'
évol 'Sambucus ebulus'	*grémola* 'moaning'	*tódol* 'astragalus'
feréstec 'wild'	*llécol* 'slate'	*tómbola* 'tombola'
focus 'focus'	*llémena* 'nit'	*tórtora* 'turtledove'
fofo 'spongy'	*llépol* 'having a sweet tooth'	*totxo* 'brick'
folro 'lining'		*tronxo* 'stalk'

Words having /ɛ/ or /ɔ/ in the stated environments are vastly more numerous than these in (34), in all dialects, though there are a few more words in which some Valencian speakers have /e/, for example, *Domènec* (surname), *època* 'age', *pèsol* 'pea', *tèrbol* 'murky'. On the other hand, Valencians may have /ɛ/ in some of the words mentioned in (34) as having /e/, such as *llécol* 'slate' or *grémola* 'moaning'. Thus, while the lexical details may vary slightly between varieties, the general principle is observed. What is the basis of the principle? Common to all words of this form are stems of at least two syllables; apart from the cases I have mentioned in which /i/ is part of a derivational suffix, all the stem syllables comprise a root. Hence there are typically more places to display phonemic contrasts than is the case in monosyllabic stems (of which there are very many in Catalan). Thus in polysyllabic stems there is less need than in monosyllabic stems to call on the full range of vocalic contrasts in order to

[23] Except when /i/ follows (see *orri* [ˈɔri] 'sheep milking parlour' above), the trill /r/ exclusively favours /o/ over /ɔ/ preceding. Hence *morro* [ˈmoru] 'snout' and its rhymes are not mentioned here.

maintain lexical distinctions. (There is also less need to call on the full range of consonantal contrasts: in a ˈ$(C_1)VC_2VC_3$ stem template C_3 is never an alveolo-palatal or palatal consonant.) Furthermore, given a tendency in Catalan, if no more than that, towards stress timing, a bisyllabic paroxytone stem, such as *geni* [ˈʒɛni] 'genius' will be inclined to approximate in length a monosyllabic stem, such as *gen* [ˈʒɛn] 'gene'. So in the first, stressed syllable of a bisyllabic stem there may be less time in which to assemble the articulatory precision to contrast seven (or eight) vowels, and equally, less time for a hearer to identify the cues distinguishing between seven (or eight) vowels. (The same principle plausibly lies behind the neutralizations of vowel contrasts involved in trisyllabic laxing in English.) These considerations add to those previously mentioned that favour neutralization of mid-vowel height contrast specifically before /i/ in a following syllable. In the more general context, that is, in paroxytones and proparoxytones before rounded vowels /u/ and /o/, or before any vowel followed by a further consonant in the stem, I have yet to account for why it is the contrast between the mid vowel pairs that is neutralized (or nearly so),[24] with the half-open vowel as the product of neutralization, in each case.

Part of the explanation for why it is the pairs of mid vowels that merge comes from typological patterns. The pairs that merge are those where a distinction is added when expanding a typical five-vowel system /i, e, a, o, u/ to a seven-vowel system /i, e, ɛ, a, ɔ, o, u/. The 'anchor' points of the system, /i, a, u/, resist neutralization relatively more readily. Similarly, when neutralization affects a five-vowel system /i, e, a, o, u/ it is the 'extra' mid vowels, /e, o/, that are lost, though in this case merger with one or more 'anchor' points, /i, a, u/, is inevitable. This is what we see in the system of unstressed syllables in north Catalan, where /e/ and /a/ of the stressed system merge (as [ə], the low non-round vowel of the unstressed system), and /o/ and /u/ of the stressed system merge as [u]. The principle behind the typology of vowel qualities in systems of different sizes is dispersion of contrast (Padgett 2001; Ní Chiosáin & Padgett 2001).

The other reason why it is the mid vowel pairs that neutralize in Catalan in certain contexts is that it is in these pairs, /e/ – /ɛ/ and /o/ – /ɔ/, that lexical phonemic contrasts are generally least stable, though this relative instability is itself an aspect of the typology of dispersion contrast mentioned in the previous paragraph. This instability of lexical mid-vowel incidence has been commented on by several Catalan linguists. I turn to examine it more closely in the next section.

2.2.1.3 *Instability in the lexical incidence of half-close and half-open mid vowels*

I have already mentioned one major source of 'instability' in the case of the mid front vowels: major dialect difference. Setting aside north Catalan, which has no

[24] Even where the height of a mid vowel is not predictable in these contexts, there are apparently no true minimal pairs, e.g. no *[ˈʎɛmənə] contrasting with *llémena* 'nit', an observation which is consistent with the interpretation of the distribution gap in terms of restricted contrastibility.

mid vowel height contrasts, we find that while there is a set of words in which all dialects have /e/, such as *déu* 'god' or *llebre* 'hare'; and there is a second set where they all have /ɛ/, including *mel* 'honey', *set* 'seven', *divendres* 'Friday', and *paciència* 'patience'; there is also a third set where the vowel quality differs between dialects, including *net* 'clean', *conèixer* 'know', *francès* 'French' (and see (29) above). But this is not the end of the 'instability'. I have already mentioned, in passing, examples of variation in vowel quality within the dialects. This kind of variation, involving both front and back mid vowel pairs, has received a good deal of attention from Catalan linguists, especially as it affects the central variety. A series of papers by Badia i Margarit (reprinted in Badia i Margarit 1988) reviewed the extent of this phenomenon. Badia i Margarit (1970 [1988]) refers to some 300 lexical items where the choice between higher and lower mid vowel is variable in Barcelona, such as *primavera* 'spring', *concentra* 'concentrate.3SG.PRS.IND', *sostre* 'ceiling', *adopta* 'adopt.3SG.PRS.IND'. Though Badia's focus in the 1970 paper is Barcelona pronunciation, he makes clear that such variation is quite extensive in the central dialect region.

Given this degree of intra-dialect variation in the incidence of /e/ versus /ɛ/ and of /o/ versus /ɔ/, Pi-Mallarach (1997) sought to investigate whether there might be tendencies currently at work towards distributing the two members of each pair on the basis of phonological context; that is, whether there was a current tendency towards complementary distribution. He investigated the pronunciation of 377 words (195 with orthographic <*e*> and 182 with <*o*>) by forty-four students in the city of Vic who were born between 1974 and 1976. The words (described as 'neologisms', and including borrowings and acronyms) were selected to be relatively unfamiliar to the informants, and to illustrate a comprehensive range of phonological environments, determined by the segment following the vowel in question, by whether the vowel in question was in an initial or a non-initial syllable, and including cases where the following syllable contained [u], given the previous observation (mentioned above) that half-open mid vowels are strongly preferred in paroxytones and proparoxytones before rounded vowels. (The environment before [i] in a following syllable was not investigated by Pi-Mallarach, on the grounds that the half-open member of each vowel pair is virtually categorical in that environment, as I have shown.) The possible outcomes in each environment investigated were (a) determined or non-contrastive distribution or (b) non-determined distribution, where a contrast was available in principle. In the latter case the distribution might be (i) non-variable (indicating secure lexical/phonemic contrast) or (ii) variable, showing greater or lesser tendency to prefer one or the other vowel. Within this particular corpus (not designed to be representative of the incidence of <*e*> and <*o*> in the vocabulary as a whole) the proportion of (b.i) type items, i.e., items showing no significant variation where a phonemic contrast was theoretically available, proved to be quite small. Not surprisingly, the (b.i) list included items such as were more likely to be established in the spoken vocabulary of students aged 18–20, such as: *bebè* [bəˈβɛ] 'baby', *calé* [kəˈle] 'dosh', *clixé* [kliˈʃe] 'photographic plate, cliché',

cromosoma [krumuˈzomə] 'chromosome', *diploma* [diˈplomə] 'diploma', *gemma* [ˈʒemmə] 'gem', *soroll* [suˈroʎ] 'noise'.

Pi-Mallarach's study confirmed observations in other sources that the half-open vowels make up the default category among the mid vowels in Catalan in the more recent part of the vocabulary: for orthographic <*e*> he observed 67 per cent [ɛ] realizations; for <*o*> he observed 61 per cent [ɔ] realizations. These proportions result from the majority of contexts categorically requiring, or markedly favouring, the more open members [ɛ], [ɔ], of the pairs. The effects of context are examined below (Pi-Mallarach 1997).

(1) Half-open [ɛ] and [ɔ] are virtually categorical in proparoxytones and consonant-final paroxytones (/_CVC(V)), such as *fèretre* [ɛ] 'bier', *fórmula* [ɔ] 'formula', *fenomen* [ɔ] 'phenomenon'. In only three of such items studied are elements of the context strong enough to shift the preference in the opposite direction: *semen* 'semen' (only 41 per cent [ɛ] in the first syllable; following [m] in an initial syllable tips the balance towards [e][25]), *pòmex* 'pumice' (48 per cent [ɔ]; preceding labial and following [m] favour [o]), and *bòrees* 'Boreas' (only 27 per cent [ɔ]; following [rV] and a preceding labial favour [o]). As the figures show, even here, a significant proportion of informants favoured the half-open vowels.

(2) Half-open [ɛ] and [ɔ] are virtually categorical also before [u] in a following syllable (see above §2.1.2). Exceptions are *quimono* [kiˈmonu] 'kimono', *biombo* 'folding screen', and *cromo* 'colour print', where following [m], [n], promote categorical [o] (see below) and *còndor* [ˈkɔndur] where [ɔ] makes only 64 per cent. In the latter case the consonant-final paroxytone environment (favouring [ɔ]) outranks the following [n] (favouring [o]). That is to say, in constraint ranking terms, and ignoring variation, *oC_1VC » *ɔ[+nasal] » *oC_1u. A following palatal lateral [ʎ] promotes a degree of raising of [ɛ] even in this environment: *colomello* 'eye tooth' (only 41 per cent [ɛ]), *tomello* 'thyme' (dialect word) (only 32 per cent [ɛ]), though following [ʎ] strongly favours [e] elsewhere.

(3) Following labial obstruents strongly favour half-open [ɛ] and [ɔ]. *Adepte* 'adept', however, only achieves 41 per cent [ɛ]. As Pi-Mallarach (125) explains, this is likely to be due to the opposing effect of lexical rhyme analogy with well-established words like *concepte* [e] 'concept' and *excepte* [e] 'except'. *Llefre* 'grimy' (see §2.2.1.1) records only 25 per cent [ɛ], and *gofre* 'waffle' only 43 per cent [ɔ]. In *llefre* the raising effect of the initial palatal lateral [ʎ] is competing with the lowering effect of following [f]. These data for *llefre* and *gofre* are also consistent with the idea that a following onset consisting of a labial obstruent + [r] has a less powerful opening effect than other following labials: compare *orfebre* 'goldsmith' with only 68 per cent [ɛ], and *cobra* 'cobra' with only 68 per cent [ɔ]. Open /ɔ/ is also some way from categorical before [f] in the first syllable of bisyllabic words (as opposed to virtually categorical [ɔ] in longer

[25] i.e. ignoring variation, *#Cɛm » *eC_1VC.

words, such as *butllofa* 'blister', *patofa* 'liar'): *mofa* 'jibe' (52 per cent [ɔ]), *tofa* 'tuft' (59 per cent [ɔ]), *plofa* 'soggy' (73 per cent [ɔ]) and *fofo* 'flabby' (only 34 per cent [ɔ], despite the lowering effect of [u] in a following syllable). It is plausible that preceding labials contribute to raising here, though preceding [l] is a countervailing force.

(4) Following dentals [t], [ð], provoke virtually categorical half-open vowels, despite the fact that /et/ and /eð/ are attested with reasonable frequency in the lexicon.

(5) Before the alveolar fricatives [s], [z] the half-open vowels are generally favoured. However, [o] is preferred in *glucosa* 'glucose', no doubt by rhyme analogy with the frequent feminine adjective suffix *-osa* /+ˈoz+a/. Before [st], /e/ and /ɛ/ appear to be in unstable contrast. In ascending order of [ɛ] variants we have the items *pedestre* 'pedestrian' (34 per cent [ɛ]), *est* 'east' (48 per cent [ɛ]), *eqüestre* 'equestrian' (55 per cent [ɛ]), *palestra* 'palaestra' (68 per cent [ɛ]), *estre* 'oestrus' (86 per cent [ɛ]), *trimestre* 'term' (95 per cent [ɛ]). This variation reflects the competing analogies of /-estrə/ and /-ɛstrə/, both of which are attested elsewhere in the lexicon. (For example, Bruguera (1990) gives *alpestre* 'alpine' with [e], but *terrestre* 'terrestrial' with [ɛ].) However, [e] might have been expected more often in *est*, since final /-ɛst/ seems to be absent from the rest of the lexicon (twenty-five rhymes in /-est/ in Rafart 1999).

(6) The half-open vowels are strongly preferred before semivowels [j] and [w], which is in line with the dissimilatory effects of [i] and [u] in following syllables. In the lexicon as a whole /ej/ is attested in *llei* 'law', *rei* 'king', and *reina* 'queen', though some speakers have /ɛ/ in those words; and/oj/ is attested in *coi* 'hammock' and *cofoi* 'smug'. The endings /-ew/ and /-ow/ are reasonably numerous (alongside /-ɛw/ and /-ɔw/).

(7) Before alveolo-palatal sibilants [ʃ], [ʒ], [t͡ʃ], one might expect the half-close [e], [o] to be preferred by anticipatory coarticulation. Rhyme analogy would also favour /-eʃ/ and /-oʃ/. But the data show rather that [ɔ] is preferred over [o] in this context. Only *motxo* 'blunted' (70 per cent [ɔ]) shows less than 75 per cent [ɔ] (and this is in the environment /_Cu that strongly favours open vowels). The front vowel pair display unstable contrast, with [e] predominating in *bretxa* 'breach', *greja* (type of silk), and *esquetx* 'sketch', but [ɛ] predominating in *quetx* 'ketch'.

(8) Before velars [k], [ɣ], the half-open vowels are categorical in the sample, though /ek/, /eɣV/, and /ok/ are quite well attested in the lexicon as a whole.

(9) The pre-nasal environment shows up a markedly different effect in the front and the back vowel sets. Between the front pair, [ɛ] is strongly preferred before [n], and [m] (with some constraints). However in the back vowel set, before nasals, [o] is almost categorical. This is in line with the general perceptual

effects of nasalization mentioned by Pi-Mallarach (1997: 151), whereby nasalization has a lowering effect on [i] and [e] but a raising effect on [æ], [a] and [o]. Though [m] generally lowers a preceding front mid vowel to [ɛ], [em] is preferred in initial syllables, perhaps by rhyme analogy with words like *crema* 'cream', *crema* 'burn.3SG.PRS.IND', *tema* 'theme'. In fact, Pi-Mallarach's set of <-*em*-> words allows us to rank some of the constraints favouring or disfavouring [ɛ]. Ignoring variation (significant numbers of [ɛ] for *lema* and *semen*), he observes *fèmur* [ɛ] 'femur', *semen* [e] 'semen', *lema* [e] 'slogan', *dilema* [ɛ] 'dilemma', *barem* [ɛ] 'schedule'. The constraint ranking is thus *eCu » *#Cɛm » *eCVC, *em, as in (35).

(35)

	*eCu	*#Cɛm	*eCVC	*em
ˈfemuɾ	*!		*	*
☞ ˈfɛmuɾ		*		
☞ ˈsemən			*	
ˈsɛmən		*!		
☞ ˈlemə				*
ˈlɛmə		*!		
diˈlemə				*!
☞ diˈlɛmə				
bəˈɾem				*!
☞ bəˈɾɛm				

Before the alveolar nasal [n], [ɛ] is favoured with little variation by Pi-Mallarach's informants, despite /en/ being quite well supported in the lexicon.

As just mentioned, half-close [o] is strongly favoured before nasals in Pi-Mallarach's study. Only the opening effect of the _CVC environment can outweigh it (hence *fenomen* 'phenomenon', *abdomen* 'abdomen', *còndor* 'condor' with majority [ɔ]); In *pòmex* 'pumice' [ɔ] reaches only 48 per cent despite the _CVC environment. A preceding labial in the initial syllable seems to be the additional factor favouring [o]. The same factor can be seen at work in *pol·len* 'pollen' with only 64 per cent [ɔ]. A following [u] has negligible countervailing effect to the raising effect of the nasal, for example *quimono* 'kimono', *cromo* 'colour print'. It does not seem realistic, however, to model the constraints on [o] and [ɔ] that we see here in a strictly ranked constraint hierarchy. It is not even evident that the various effects are simply additive.

(10) The environment before an onset alveolar tap (_ɾV) can probably best be interpreted as permitting phonemic height contrast. In words of more than two syllables [-eɾə] is strongly preferred over [-ɛɾə] (e.g. in *estratosfera* 'stratosphere'), though this is plausibly attributable to rhyme analogy with the frequent *-era* suffix. In shorter *xera* 'spree', 'endearment', [ɛ] was preferred by 61 per cent. Half-close [o] in *angora* 'angora', etc. is supported by analogy with the suffix *-ora*, but [ɔ] is favoured (68 per cent) in *folklore* [ful'klɔɾə] 'folklore'.

Before the alveolar trill [r], both intervocalic and pre-consonantal (underlying /R/) [ɛ] is categorical in Pi-Mallarach's investigation, not only before intervocalic [r] and [r] followed by [+coronal] consonants, where [ɛ] is categorical in the general lexicon (see §2.2.1.1 above), but before [r] followed by non-coronal consonants as well, as in *iceberg* 'iceberg' and *paquiderm* 'pachyderm'. Likewise [ɔ] is favoured before [r] followed by a consonant (e.g. *cloroform* [ɔ] 'chloroform') despite the presence of /-orC/ in the lexicon; only 32 per cent [ɔ] for *corm* 'corm', and 59 per cent for *horda* 'horde', points to unstable contrast here, though. But before intervocalic [r] only [o] occurs, in *cotorra* 'parrot' and *masmorra* 'dungeon', in line with the rest of the lexicon. (*Pròrroga* 'deferment' has [ɔ] not because it is followed by [u] in the next syllable, as Pi-Mallarach claims (p. 147), but because it is followed by CVC. *Morro* ['moru] 'snout' shows that following [u] is insufficient to lower [or] to [ɔr].) Despite Pi-Mallarach's explanation for general anticipatory lowering of the body of the tongue in anticipation of [r], further investigation seems needed into why [e] in particular seems incompatible with following [r] in Catalan, and why [-ɔrə] is missing when, for example, [-ɔrC] is allowed, and even preferred.

(11) As Pi-Mallarach (1997: 143) points out, the predorsal retraction characteristic of the velarized alveolar lateral [ɫ] is more compatible with [ɛ], [ɔ] than with [e], [o], given that the lower mid vowels also involve a degree of pre-dorsal retraction. And, indeed, in the 'neologisms' [ɛ] is virtually categorical before [ɫ] (as it is in the lexicon as a whole), and [ɔ] is strongly preferred there, despite support from /ol/ elsewhere in the lexicon (e.g. in *gola* 'throat'). Only in *pol·len* 'pollen' (64 per cent [ɔ]) does the score for [ɔ] fall below 75 per cent—the preceding word-initial labial may be the 'raising' influence here; in none of the other <*-ol-*> neologisms is <*o*> preceded by an initial labial.

(12) The following palatal lateral [ʎ] provides the context most strongly predisposing towards the half-close mid vowels [e], [o], which anticipate the close tongue position of [ʎ]. This is more marked in the case of the front [e] where the articulatory overlap is closer. Despite a minority presence of /ɛʎ/ in the rest of the lexicon, [eʎ] is as good as categorical in the neologisms, except in *colomello* 'eye-tooth' (41 per cent [ɛ]), *tomello* 'thyme' (dialect word) (32 per cent [ɛ]), where following [u] is the lowering influence. Half-open [ɔ] before [ʎ] makes a greater showing, though only in certain words, suggesting unstable contrast of /o/ vs. /ɔ/ in this context (supported by the presence of both /oʎ/ and /ɔʎ/ in the lexicon). In *crioll* 'creole' [ɔ] reaches 57 per cent, a fact which Pi-

Mallarach (p. 160) attributes to the dissimilatory effect of the preceding [i] (though a comparable effect is not observed in *biombo* 'folding screen', which has [o] before a nasal, despite the preceding [i]/[j] and the potential lowering effect of following [u]). As Pi-Mallarach notes, despite very similar tongue articulation, [j] promotes dissimilation in preceding mid vowels, while [ʎ] promotes assimilation. This is plausibly because assimilation of vowel to [j] in -Vj- sequences endangers the identifiability of [j]: [ej] is harder to distinguish perceptually from [e] than [ɛj] is. The laterality of [ʎ] is sufficient to prevent perceptual fusion with preceding vowels of whatever quality. The stronger closing effect of [ʎ] than of the sibilant alveolo-palatals [ʃ], [t͡ʃ], and [ʒ] (see above) may be due to [ʎ] being articulatorily fronter than the sibilant alveolo-palatals.

(13) In prevocalic position (essentially before [ə], since hiatus before a high vowel is extremely unusual: see §3.2.3), half-close vowels are favoured, a phenomenon which is attributable to dissimilation (avoiding perceptual fusion) since [ə] tends to be half-open. In this context a potential phonemic contrast is available (Rafart (1999) records *gelea* 'jelly' with [ɛ] and *boa* 'boa' with [ɔ], alongside a majority with [-eə] and [-oə], though Pi-Mallarach's informants only offered 11 per cent [ɔ] in *boa*). Pi-Mallarach (1997: 163) gives a number of reasons to suppose that a linguistic change is in progress, with [-eə] replacing [-ɛə]. At the beginning of the twentieth century [-ɛə] was noted as the predominant pronunciation in words like *odissea* 'odyssey'; towards the end of this century pronunciation dictionaries differed in the pronunciation given for words like *aldea* 'hamlet', *assemblea* 'assembly', *azalea* 'azalea', and *brea* 'tar'; in a pilot study Pi-Mallarach's informants displayed previously unobserved vowel alternation in masculine ([-ɛw]) and feminine ([-eə]) forms of adjectives such as *arameu*/*aramea* 'Aramaic', *egeu*/*egea* 'Aegean', *europeu*/*europea* 'European', *maniqueu*/*maniquea* 'Manichaean', *plebeu*/*plebea* 'plebeian'; and the variation in pronunciation of his <-*ea*> neologisms was in part attributable to his informants falling into two groups, one of which favoured [-eə] and the other [-ɛə].

(14) Absolute final position revealed a contrast, though not a completely stable one, between /e/ and /ɛ/. (The back vowels were not investigated in this context.) Of fourteen words investigated, *bebè* 'baby' and *comitè* showed a clear majority for /ɛ/, *canapè* 'canapé', *consomé* 'consommé', and *ximpanzè* 'chimpanzee' showed the two vowels equally divided. The remaining words showed preference for /e/, categorically in five cases.

What we may conclude from Pi-Mallarach's investigation is that in central Catalan the potential for lexical contrast using /e/ vs. /ɛ/, and /o/ versus /ɔ/, is far from fully exploited in words that fall into the class of what Pi-Mallarach calls 'neologisms'. Rather, the tendency is to distribute the vowels in accordance with phonetic context, or by analogy with tendencies already present in the lexicon. This is perhaps not altogether surprising, since many such words may be acquired primarily via the written media, and Catalan orthography distinguishes the members of the pairs relatively rarely; or the words may become familiar via

their Spanish pronunciation in the spoken mass media, and Spanish lacks a phonemic distinction between half-close and half-open vowels. However, the effects of phonological context in favouring one or other alternative prove quite interesting, insofar as they are phonetically motivated, or can plausibly be argued to be so, without being phonetically determined—they could be otherwise than they are (in part at least) without violating phonetic naturalness. In this area, then, we see orderly and principled choices made by particular speech communities from among the set of 'natural' possibilities. In some cases it is possible to model these choices in a constraint hierarchy, though in general variation of the kind seen here might be more realistically modelled by weighted constraints (see Boersma 1998).

2.3 VOWEL REDUCTION

2.3.1 *Introduction*

Vowel reduction—the neutralization of some vowel contrasts in unstressed syllables—is a feature of all Catalan varieties. As lexical words have only one stress, and as many suffixes bear lexical stress, the effects of vowel reduction are widely attested in inflected and derived words. However, the dialects display several different patterns of vowel reduction, some of them more extensive than others. This section starts with a brief descriptive summary of the vowel reduction patterns. In subsequent sections I consider how best to account for these patterns from an OT perspective, and then examine some exceptions to vowel reduction.

2.3.2 *Vowel reduction in western Catalan*

As mentioned in §2.1, western Catalan dialects have a seven-vowel system in stressed syllables, with a height contrast among front and back mid vowels as in (36).

(36)

i	u
e	o
ɛ	ɔ
a	

In unstressed syllables half-open /ɛ/ and /ɔ/ are replaced by /e/ and /o/ respectively, resulting in neutralization, as can be seen in the examples in (37) representing standard Valencian pronunciation.

(37) Basic pattern of western Catalan vowel reduction (Valencian)

	Base form	Base + stressed suffix
(a)	*gel* [ˈd͡ʒɛl] 'ice'	*gelat* [d͡ʒeˈlat] 'ice cream'
	perdre [ˈpɛɾðɾe] 'lose.INF'	*perdem* [peɾˈðem] 'lose.1PL.PRS.IND'
	pera [ˈpeɾa] 'pear'	*perera* [peˈɾeɾa] 'pear tree'
(b)	*cosa* [ˈkɔza] 'thing'	*coseta* [koˈzeta] 'thing.DIM'
	porta [ˈpɔɾta] 'carry.3SG.PRS.IND'	*portem* [poɾˈtem] 'carry.1PL.PRS.IND'
	tot [ˈtot] 'all.M.SG'	*total* [toˈtal] 'total'

This pattern of vowel reduction, from a seven-vowel stressed system with half-open vowels to a five-vowel system without height contrasts in mid vowels, is found also in Portuguese and Italian, and seems to have been an early common feature of proto-Romance. The other Catalan dialects show at least this degree of vowel reduction in all unstressed syllables, though they add further reductions. (The height contrast in mid vowels derives from a quantity contrast in Latin, and late spoken Latin appears to have lost quantity distinctions in unstressed syllables while replacing them with quality distinctions in stressed syllables.) This five-vowel system will be the main focus in this section. It is of some interest, though, that there are, additionally, significant restrictions on the contrast between /a/ and /e/ in unstressed syllables in western Catalan. In post-tonic syllables /a/ and /e/ contrast only in final open syllables; in closed or non-final syllables only /e/ is found. Thus, for example, the plural of both *pobre* [ˈpɔβɾe] 'poor.M' and *pobra* [ˈpɔβɾa] 'poor.F' is *pobres* [ˈpɔβɾes] 'poor.PL'. Something like *[ˈpɔβɾas] is ill-formed in all dialects of Catalan, except in Ribagorça (north-eastern Aragon). And in onsetless initial closed syllables, at least in informal speech in western Catalan, /a/ is found to the exclusion of /e/, so *angoixa* [aŋˈgojʃa] and *entenc* [anˈteŋk] both begin with the same vowel. In western Catalan the realization of word-final /a/ is subject to considerable variation, including, in some Valencian varieties, interesting patterns of vowel harmony agreeing with stressed half-open vowels, giving, for example, *tela* [ˈtɛlɛ] 'cloth', *cosa* [ˈkɔzɔ] 'thing'. Valencian vowel harmony has been investigated from an OT perspective by Jiménez (1998; 2001).

2.3.3 *Vowel reduction in Majorcan*

Balearic Catalan, it will be recalled, displays a system of eight vowels in stressed syllables, as in (38).

(38)

i		u
e		o
	ə	
ɛ		ɔ
	a	

As in western Catalan, the half-open mid vowels are absent from unstressed syllables in Majorcan, so unstressed /ɔ/ becomes [o].[26] But all the non-high unrounded vowels merge as schwa [ə] in unstressed syllables, giving a basic unstressed four-vowel system, /i, ə, o, u/. However, unstressed [e] is found also in Majorcan, variably, in certain circumstances, in correspondence with underlying /ˈe/ when alternations are available, giving a five-vowel /i, e, ə, o, u/ unstressed system. The presence of unstressed [e] is favoured by some combination of morphological alternation, favourable phonological context, and interference from Spanish (Bibiloni 1998; §2.3.12 below). Setting this deviant behaviour of /e/ aside for the moment, the examples of (37) (with additions) appear in Majorcan as in (39).

(39) Basic pattern of Majorcan vowel reduction

	Base form	Base + stressed suffix
(a)	*gel* [ˈʒɛl] 'ice'	*gelat* [ʒəˈlat] 'ice cream'
	perdre [ˈpɛrðɾə] 'lose.INF'	*perdem* [pərˈðəm] 'lose.1PL.PRS.IND'
	pera [ˈpərə] 'pear'	*perera* [pəˈɾerə] 'pear tree'
	pedra [ˈpeðɾə] 'stone'	*pedrera* [pəˈðɾerə] 'quarry'
	banya [ˈbaɲə] 'bathe.3SG.PRS.IND'	*banyam* [bəˈɲam] 'bathe.1PL.PRS.IND'
(b)	*cosa* [ˈkɔzə] 'thing'	*coseta* [koˈzətə] 'thing.DIM'
	porta [ˈpɔɾtə] 'carry.3SG.PRS.IND'	*portam* [poɾˈtam] 'carry.1PL.PRS.IND'
	tot [ˈtot] 'all.M.SG'	*total* [toˈtal] 'total'

2.3.4 *Vowel reduction in eastern Catalan, except Majorcan*

The remaining eastern Catalan varieties, i.e. central and north Catalan and the dialects of Minorca and Ibiza, display the most radical pattern of vowel reduction, such that unstressed syllables have generally only [i, ə, u], with some cases of [e] and [o] in limited contexts. So, as in Majorcan, non-high non-rounded vowels are neutralized as [ə] in unstressed syllables, and in addition all rounded vowels are neutralized as [u]. In all Catalan varieties underlying /i/ and /u/ are unaffected by vowel reduction; the realization of these phonemes in unstressed positions does not differ significantly from their realization in stressed positions (Recasens 1991b; 1993). The example words of (39) are realized in central Catalan in the form illustrated in (40).

[26] Most of Majorca also displays an assimilative process of height neutralization whereby unstressed /o/ becomes [u] before /i/ (stressed or unstressed) in the next syllable, e.g. *cobrir* [kuˈβɾi] 'cover.INF' vs. *cobreix* [koˈβɾəʃ] 'cover.3SG.PRS.IND'. But this raising process is not found in the major towns such as Palma, Manacor, and Inca, and is recessive (though very ancient) elsewhere.

(40) Basic pattern of vowel reduction in eastern Catalan, except Majorcan

	Base form	Base + stressed suffix
(a)	*gel* [ˈʒɛl] 'ice'	*gelat* [ʒəˈlat] 'ice cream'
	perdre [ˈpɛɾðɾə] 'lose.INF'	*perdem* [pəɾˈðɛm] 'lose. 1PL.PRS.IND'
	pera [ˈpɛɾə] 'pear'	*perera* [pəˈɾɛɾə] 'pear tree'
	pedra [ˈpeðɾə] 'stone'	*pedrera* [pəˈðɾɛɾə] 'quarry'
	banya [ˈbaɲə] 'bathe.3SG.PRS.IND'	*banyem* [bəˈɲɛm] 'bathe. 1PL.PRS.IND'
(b)	*cosa* [ˈkɔzə] 'thing'	*coseta* [kuˈzɛtə] 'thing.DIM'
	porta [ˈpɔrtə] 'carry.3SG.PRS.IND'	*portem* [purˈtɛm] 'carry. 1PL.PRS.IND'
	tot [ˈtot] 'all.M.SG'	*total* [tuˈtal] 'total'

The three patterns of vowel reduction that have been introduced above show increasing degrees of markedness constraints interacting with faithfulness. In relation to the western Catalan pattern of seven vowels reduced to five, Majorcan promotes a further constraint merging the remaining non-high non-round vowels, and the rest of central Catalan promotes another constraint merging the remaining rounded vowels. The most restrictive three-vowel system, [i, ə, u], is one that retains no essential height contrasts, and that can be described as realizing three vowel 'colours', Palatal ([i]), Labial ([u]) and 'neither Palatal nor Labial' ([ə]).

2.3.5 *The interpretation of Catalan vowel reduction*

The functional OT account proposed here develops the approach of Crosswhite (2004). According to Crosswhite, there are two types of phonetically motivated vowel reduction (whose effects may overlap to a degree). A 'contrast-enhancing reduction' aims to ensure that, in a context disfavouring perceptual contrast between vowels, the contrast retained should be the best available from the point of view of perception (Lindblom 1986). A particularly good contrast-enhancing reduction is one that avoids mid vowels altogether in unstressed syllables, and uses only the 'corner' vowels [i], [a], [u]. Observe that contrast-enhancing reduction favours [a] in unstressed position. The other type of phonetically motivated vowel reduction is what Crosswhite calls 'prominence reduction'. There is an analogy to be drawn with prominence within the syllable, whereby certain sound types (those of high sonority) are best suited to a position of high prominence (the nucleus or peak), while low sonority sound types are better suited to low prominence positions (margins). To follow through the analogy, unstressed nuclei stand to stressed nuclei in the same relation as margins stand to peaks. For Crosswhite (p. 207), the relevant prominence scale for stressed and unstressed vowels is low-frequency amplitude, so [a] (and other low peripheral vowels) are well suited to stressed syllables, whereas [i], [u], and

in particular [ə] are well suited to unstressed syllables. Note that by 'prominence reduction' [ə] is strongly favoured over [a] in unstressed position; high 'corner' vowels [i] and [u] are well suited to unstressed position from either perspective.

2.3.6 *Excluding* [ɛ, ɔ] *from unstressed syllables*

I first examine the unstressed five-vowel [i, e, a, o, u] system from the contrast-enhancing perspective. Crosswhite (2004: 194–7) explains why the three-vowel 'corner' system [i, a, u] is a good one perceptually. These three vowels are maximally distinctive (make best use of dispersion within the vowel space: Lindblom 1986); they are stable or robust, in that they are subject to quantal effects, whereby large changes in articulation produce rather small acoustic consequences;[27] and they are 'focal' inasmuch as they have spectral prominence caused by convergence of formants—of F_3 and F_4 in the case of [i], of F_2 and F_1 in the case of [a], and of F_1 and F_0 in the case of [u]. The perceptually good [i, a, u] system is well suited to contexts where distinctiveness is potentially compromised, such as (in the case in question) unstressed syllables, which are likely to be shorter in duration and lower in intensity. (Crosswhite suggests that vowel reduction is absent from those languages in which stress does not correlate with duration.) The licensing constraint that, Crosswhite proposes, motivates most cases of contrast-enhancing vowel reduction is as in (41).

(41) LIC-Noncorner/Stress: Non-corner vowels are licensed only in stressed positions (Crosswhite 2004: 194).

The Catalan five-vowel unstressed system does not display the effects of the full LIC-Noncorner/Stress constraint. Rather, only *some* non-corner vowels are licensed only in stressed position, namely, half-open mid vowels. The formulation of the constraints appropriate for Catalan depends to some degree on how the overall vowel system is described. I propose here a system of features (mostly unary), arranged in feature classes (Padgett 1995a; 1995b), without feature geometry:

Vowel height features (Height): High, Low, Mid [+close, −close][28]
Vowel colour features (Colour): Palatal, Labial

These features apply to Catalan vowels in the manner illustrated in (42).

(42)

		Palatal		Labial
High		i		u
Mid	+close	e	ə	o
	−close	ɛ		ɔ
Low			a	

[27] The vowels [ʌ] and [ə] are also 'quantal' in this sense, but [i, a, u] are perceptually superior in being both quantal and surrounded by areas of greater acoustic instability.

[28] Ever since the feature(s) [Advanced Tongue Root/Retracted Tongue Root] were introduced into phonological theory, these have been used in work on Catalan phonology to capture the

The active licensing constraint is then LIC[−close]/Stress and the relevant constraint hierarchy is as in (43), using feature IDENTITY as the faithfulness constraint type. (Nothing hangs on the particular version of featural faithfulness.)

(43) IDENTHigh, IDENTLow, IDENTColour, LIC[−close]/Stress » IDENT[clo]

The tableau in (44) evaluates various unstressed vowel candidates for *gelat* /ˈd͡ʒɛlˈad/ 'ice cream' (44a), *portem* /ˈpɔRtˈe+m/ 'carry.1PL.PRS.IND' (44b), and *banyem* /ˈbaɲˈe+m/ 'bathe.1PL.PRS.IND' (44c) in Valencian.

(44)

	Valencian	IDENTHigh, IDENTLow, IDENTColour	LIC[–clo]/Stress	IDENT[clo]
(a)	d͡ʒiˈlat	*!		
	☞ d͡ʒeˈlat			*
	d͡ʒɛˈlat		*!	
	d͡ʒəˈlat	*!		
(b)	puɾˈtem	*!		
	☞ poɾˈtem			*
	pɔɾˈtem		*!	
	pəɾˈtem	*!		
(c)	beˈɲem	*!		
	bəˈɲem	*!		
	☞ baˈɲem			

2.3.7 *Basic vowel reduction in eastern Catalan*

In comparison with the pattern of contrast-enhancing reduction considered in §2.3.6, the notable characteristic of vowel reduction in eastern Catalan is the presence of [ə] in unstressed syllables, whether (as in Balearic) or not (as in continental Catalan) such a vowel is part of the inventory of stressed syllable nuclei. For Crosswhite, vowel reduction involving target [ə] is characteristic of the prominence reduction strategy. As mentioned above, the understanding and

distinction between half-close and half-open vowels. But this use is merely diacritic. There is no evidence in Catalan for tongue root advancement or retraction independent of vowel height. In the text the values [±close] of the feature Mid are proposed. Other plausible alternatives could do the job just as well.

formulation of prominence reduction follows the model of sonority hierarchies in syllable structure as developed by Prince & Smolensky (1993). The relevant constraint hierarchy is calculated by crossing two phonetic prominence scales as in (45) (Crosswhite 2004: 209).

(45) Accentual prominence scale:
stressed$_{prom}$ > unstressed
Vocalic prominence scale:
a$_{prom}$ > ɛ, ɔ$_{prom}$ > e, o$_{prom}$ > i, u$_{prom}$ > ə

The phonetic correlate of vocalic prominence is suggested by Crosswhite to be low-frequency amplitude; part of this is a question of inherent length. For Catalan, Recasens (1991b: 90, 105) describes [a] as longer than other vowels, while [ə] is shorter than other vowels. Part of the prominence reduction constraint hierarchy derived from crossing the two scales in (45) is given in (46).

(46) *UNSTRESSED/a » *UNSTRESSED/ɛ, ɔ » *UNSTRESSED/e, o »
*UNSTRESSED/i, u » *UNSTRESSED/ə

In eastern Catalan the two left-hand constraints of the prominence reduction hierarchy, *UNSTRESSED/a and *UNSTRESSED/ɛ, ɔ, are undominated,[29] while the two right-hand ones, *UNSTRESSED/i, u and *UNSTRESSED/ə, are inactive. It is the interaction of *UNSTRESSED/e, o with other constraints, whether of markedness, faithfulness, or paradigm uniformity, that gives rise to some complexities and variation. I consider first here in tableau (47) the basic pattern of vowel reduction in eastern Catalan (excluding Majorcan), where *UNSTRESSED/e, o is highly ranked. I use the same lexical examples as in (44) above: *gelat* /ˈʒɛl+ˈad/ 'ice cream' (47a), *portem* /ˈpɔRt+ˈɛ+m/ 'carry.1PL.PRS.IND' (47b), and *banyem* /ˈbaɲ+ˈɛ+m/ 'bathe.1PL.PRS.IND' (47c) (central Catalan pronunciation and morphology—note that the present indicative theme vowel is /+e/ in western Catalan but /+ɛ/ in continental eastern Catalan).

[29] An isolated exception in central Catalan is the frequent pronunciation of *Nadal* 'Christmas' as [naˈðal], alongside regular [nəˈðal]. There is no satisfactory explanation. Can it result from phonetic reinforcement of the pun that appeals to children: *Nadal* = *'nar dalt* [naˈðal] 'to go up'? At any rate the pronunciation [naˈðal] is consistent with that of a compound of two roots.

(47)

	IDLab	*UNSTR/a	*UNSTR/ɛ,ɔ	*UNSTR/e,o	IDHeight	IDPal
ʒiˈlat					*!	
ʒeˈl at				*!	*	
ʒɛˈl at			*!			
ʒaˈl at		*!			*	*
☞ ʒəˈl at						*
☞ purˈtɛm					*	
porˈtɛm				*!	*	
pɔrˈtɛm			*!			
pɐrˈtɛm	*!					
beˈɲɛm				*!	*	*
☞ bəˈɲɛm					*	
baˈɲɛm		*!				

(Continental E. Catalan): IDENTLab, *UNSTR/a » *UNSTR/ɛ, ɔ » *UNSTR/e, o » IDENTHeight » IDENTPal

The constraint ranking in (47) shows IDENTLab highly placed: vowel reduction does not affect rounding; for this reason candidates with rounding violations are mostly ignored. The constraint IDENTHeight is interpreted here in such a way that [ə], which lacks a [±close] value, corresponds in height to any mid vowel, either [+close] or [−close]. Consequently the candidate [ʒəˈlat] (for input /ʒɛl+ˈad/) carries no mark under IDENTHeight, and thereby wins over *[ʒiˈlat] with a height violation. A different approach might evaluate height violation in a gradient way, such that [ʒəˈlat] scored one violation (within the Mid category), whereas *[ʒiˈlat] would score two violations (between height categories). In any case, the logic of constraints of the IDENTHeight type is that with respect to input /ʒɛl+ˈad/, *[ʒiˈlat] deviates further in height than [ʒəˈlat] does. It is the pair *[ʒiˈlat] ≺ [ʒəˈlat] also that reveals the ranking IDENTHeight » IDENTPal. Faithfulness to front-vowel quality is less important than (relative) height faithfulness. In the tableau and the associated hierarchy I have kept *UNSTR/e, o as a single constraint. However, the data only demonstrate the following rankings:

*UNSTR/e » IDENTPal: *[ʒeˈlat] ≺ [ʒəˈlat]
*UNSTR/o » IDENTHeight: *[porˈtɛm] ≺ [purˈtɛm]
IDENTHeight » IDENTPal: *[ʒiˈlat] ≺ [ʒəˈlat]

I shall have occasion to return to this detail below (§2.3.9).

2.3.8 *Basic vowel reduction in Majorcan*

As described in §2.3.3, the unstressed vowel system of Majorcan differs from the general pattern of eastern Catalan in that a contrast of [o] and [u] is maintained. (As in western Catalan, /o/ and /ɔ/ are neutralized as [o].) That is to say, the constraint *UNSTRESSED/e, o is broken up, with *UNSTRESSED/e remaining in the original position of (47) and *UNSTRESSED/o being demoted below the height faithfulness constraints. The motivation for this separation is unclear. In particular it is unclear why *UNSTRESSED/o should be demoted rather than *UNSTRESSED/e, if the problem were that [i, e, ə, o, u] is a sub-optimal unstressed vowel system (which it surely is from the perspective of contrast-enhancing reduction). Contrasts between back vowels are somewhat more marked typologically than between front vowels; that is, a height contrast in back vowels normally entails the same or more contrasts in front vowels (Ladefoged & Maddieson 1996). Demotion of *UNSTRESSED/o also entails separating out elements of the IDENTHeight family, for reasons similar to those leading to this conclusion in western Catalan (44). Specifically, IDENTHigh, IDENTLow outrank IDENT[clo]. The relevant ranking, illustrating *portam* /ˈpɔRt+ˈa+m/ 'carry.1PL.PRS.IND' in Majorcan, is given in tableau (48).

(48)

Majorcan	*UNSTR/ɛ,ɔ	*UNSTR/e,	IDENTHigh, IDENTLow	IDENT[clo]	IDENTPal	*UNSTR/o
purˈtam			*!			
☞ porˈtam				*		*
pɔrˈtam	*!					

It is interesting to consider the consequence of joint demotion of both parts of *UNSTRESSED/e, o in Majorcan to the same position, that is:

*UNSTR/ɛ, ɔ » IDENTHigh, IDENTLow » IDENT[clo] » IDENTPal » *UNSTR/e, o

Underlying /ɛ/ would surface as [ə] in unstressed syllables, just as in *gelat* in (47); *[ʒeˈlat] being knocked out due to its violating IDENT[clo]. However, underlying /e/ would surface as [e], provided IDENTPal outranks *UNSTR/e, o. This pattern is illustrated in tableau (49), where one can compare the evaluation of *gelat* with that of *pedreta* /ˈpedR+ˈət+a/ 'stone.DIM'. This outcome is, in fact, appropriate for these particular words. However, underlying /e/ surfaces as [e] in unstressed syllables only in restricted circumstances. The same root /pedR/ with a different suffix /+ˈeR+a/ gives *pedrera* [pəˈðrɾeɾə] 'quarry', showing reduction to [ə] in the unstressed syllable. I review the ranking IDENTPal » *UNSTR/e, o later when these circumstances are looked at in more detail (§2.3.12).

(49)

	Majorcan *gelat* /'ʒɛl+'ad/	*UNSTR/ɛ, ɔ	IDHigh, IDLow	ID[clo]	IDPal	*UNSTR/e, o
(a)	ʒi'lat		*!			
	ʒe'lat			*!		*
	ʒɛ'lat	*!				
	ʒa'lat		*!		*	
	☞ ʒə'lat				*	
(b)	*pedreta* /'pedR+'ət+ə/					
	pi'ðrətə		*!			
	☞ pe'ðrətə					*
	pɛ'ðrətə	*!		*		
	pa'ðrətə		*!		*	
	pə'ðrətə				*!	

Observe that while western Catalan and Majorcan both admit [e] and [o] in unstressed syllables, in western Catalan both /e/ and /ɛ/ surface as [e], because all other vowel quality identity constraints dominate IDENT[clo]. In Majorcan, however, IDENT[clo] dominates IDENTPal, and only underlying /e/ surfaces (sometimes) as unstressed [e]. For underlying /ɛ/, the IDENT[clo] violation is fatal to [e], and [ə] is the best alternative.

2.3.9 *Exceptions to eastern Catalan vowel reduction, 1:* *[əˈa], *[ˈaə], *[əə]

Throughout eastern Catalan, vowel reduction does not centralize a mid front vowel if the result would be a sequence of two 'colourless'[30] vowels: *[əa], *[ˈaə], *[əə]. Some examples are given in (50).

(50) *idea* [iˈðeə] 'idea' — *ideal* [iðeˈal] 'ideal' *[iðəˈal]
idealisme [iðeəˈlizmə] 'idealism' *[iðəəˈlizmə]
crea [ˈkɾeə] 'create.3SG.PRS.IND' — *crear* [kɾeˈa] 'create.INF' *[kɾəˈa]
creador [kɾeəˈðo] 'creator' *[kɾəəˈðo]
aeri [əˈɛɾi] 'of air' — *aeròbic* [əeˈɾɔβik] 'aerobic' *[əəˈɾɔβik]

[30] Or [+back, −round] in traditional features. An alternative description would say that [ə] and [a] are both A-coloured (in contrast to palatal and labial vowels), as in Ch. 3 here.

Israel [irəˈɛl] ‘Israel’	*israelià* [irəeˈlja] ‘Israeli’ *[irəəˈlja]
	àrea [ˈareə] ‘area’ *[ˈarəə]
	Aglae [əˈɣlae] *[əˈɣlaə]
	Dànae [ˈdanəe] *[ˈdanəə]
	reanalitza [reənəˈlidd͡zə] ‘reanalyse.3SG.PRS.IND’ *[rəənə-]

There is some evidence that the sequences [əˈa], [ˈaə], [əə] are not ruled out absolutely. See the examples in (51).

(51)	*Baal* [bəˈal]	*vehement* [bəəˈmen] ‘vehement’
	Abraham [əβɾəˈam]	*aprehensió* [əpɾəənˈsjo] ‘perception’
	Sahara [səˈarə] or *Sàhara* [ˈsaəɾə] ‘Sahara’	*neerlandès* [nəərlənˈdɛs] ‘Netherlandic’
		Aaron [əəˈɾon]
	saharià [səəˈɾja] ‘Saharan’	

The crucial difference between the examples of (50) and those of (51) is that the former have an underlying mid front vowel that is not reduced when a stressed suffix is added. In the examples of (51) neither of the adjacent vowels is underlyingly front. Perhaps not surprisingly, there are some words that vary between the pattern of (50) and (51). Recasens (1993: 93) mentions *balneari* ‘spa’, *preàmbul* ‘preamble’, *enneàgon* ‘nonagon’, and *genealogia* ‘genealogy’. Alternative lexical representations seem to offer the most likely explanation for this variation. Variation is more surprising in *paellada* [pəəˈʎaðə] ∼ [pəeˈʎaðə] ‘frying-panful’ (*paella* [pəˈeʎə] ‘frying pan’ /+ˈada/) and *graellada* [gɾəəˈʎaðə] ∼ [gɾəeˈʎaðə] ‘mixed grill’ (*graella* [gɾəˈeʎə] ‘grill’ /+ˈada/) (Recasens 1993: 93; Badia i Cardús 2001: 131) where a lexical front vowel is obvious enough in the base. Presumably those speakers who do use [əə] in these words have lexicalized them, breaking the morphological link.

The preference for [eˈa], [ˈae], [eə], [əe] over *[əˈa], *[ˈaə], *[əə] is plausibly attributed to ‘dissimilation’ (though, given the point made in the previous paragraph, ‘non-assimilation’ would be a better term; Palmada (1994a: 45) deals only with *[əə] in terms of OCP violation). To be more precise, the elements of [əˈa], [ˈaə], [əə] are not sufficiently dissimilar from one another to reliably avoid fusion, at least in perception. Remember that [ə] is the shortest vowel and [a] the longest, so [əˈa], [ˈaə] are both at risk of being heard as [ˈa], and [əə] of being heard as [ə] (in fact, [ə] is a real variant outcome of [əə] in words like *aprehensió* (51)). In terms of vowel quality, [ə] and [a] differ only in height, and this difference between mid and low, in non-labial and non-palatal vowels, is not contrastive anyway in central Catalan, where [ə] and [a] are the surface allophones of a lexical vowel without distinctive specification for height. When vowels of these qualities are adjacent in word sandhi contexts (§§4.2–4.4), the sequence is in fact reduced by deletion of [ə], as in (52).

(52) *aquesta àrea* [əˈkɛsˈtareə] 'this area'
pla africà [ˈplafriˈka] 'African plan'
mateixa amplària [məˈteʃəmˈplaɾjə] 'same width'

Thus the OCP-type constraint spelt out in (53) is well motivated. I subsequently use the abbreviation *OCP/V_A for convenience. In (53) the *OCP/V_A constraint is followed by the vowel reduction constraint hierarchy for eastern Catalan (except Majorcan), repeated from (47).

(53) *$V_{-Colour}$.$V_{-Colour}$: Colourless vowels do not stand adjacent to one another
Equivalently:
*OCP/V_A: A-coloured vowels do not stand adjacent to one another

VR Constraint hierarchy (eastern continental Catalan): IDENTLab, *UNSTR/a » *UNSTR/ɛ, ɔ » *UNSTR/e, o » IDENTHeight » IDENTPal

However, *OCP/V_A cannot simply dominate *UNSTR/e in the VR constraint hierarchy, or else [eə]/[əe] would always be preferred to [əə], whatever the input. Rather, *OCP/V_A and *UNSTR/e are equal in ranking. They cannot choose between candidates that violate each of them, so the evaluation passes down the hierarchy, to IDENTPal. To put it another way, the grammar of eastern Catalan finds the OCP violation and the presence of [e] in an unstressed syllable equally unsatisfactory. In such circumstances, the candidate that is more faithful to the input is the one that is less bad. However, the ranking in the VR constraint hierarchy in (53) which I established provisionally in §2.3.7, would, for *ideal*, allow a candidate like *[iðiˈal] to win, as its IDENTHeight violation is ranked lower than *OCP/V_A, which penalizes *[iðəˈal], and lower than *UNSTR/e, which penalizes [iðeˈal]. This consequence shows that the *UNSTR/e and *UNSTR/o constraints must be separated. *IDENTHeight is subordinate to *UNSTR/o ([purˈtɛm] ≻ *[porˈtɛm]), but outranks *UNSTR/e ([iðeˈal] ≻ *[iðiˈal]). Some relevant candidates are evaluated by this modified hierarchy in (54). Candidates violating undominated *UNSTR/a » *UNSTR/ɛ, ɔ are not considered.

An alternative approach might be to invoke a contextual faithfulness constraint dominating *UNSTR/ə, as in (55).

(54)

		*UNSTR/o	IDENT Height	*OCP/VA	*UNSTR/e	IDPal
(a)	*ideal* /i'de+'al/					
	☞ iðe'al				*	
	i ðə'al			*		*!
	iði'al		*!			
(b)	*idealisme* /i'de+'al+'iz+m/					
	i ðəə'lizmə			*		*!
	☞ iðeə'lizmə				*	
	iðəe'lizmə				*	*!*
	iðiə'lizmə		*!			
(c)	*vehement* /bAA'ment/					
	☞ bəə'men			*		
	beə'men				*	*!
	bəe'men				*	*!
	bəi'men		*!			*
	biə'men		*!			*
(d)	*portem* /'pɔRt+'ɛ+m/					
	por'tɛm	*!				
	☞ pur'tɛm		*			

E Catalan VR constraint hierarchy: *UNSTR/o » IDENTHeight » *OCP/VA, *UNSTR/e » IDENTPal

(55) IDENTPalatal/A: Adjacent to an A-coloured vowel, an input Palatal vowel remains Palatal in the output.

The issue is essentially whether the avoidance of *[əə], and similar sequences, is taken to be a matter of markedness, as a result of *OCP/V_A, or of contextual faithfulness (dominating vowel reduction), as with IDENTPalatal/A. The tableau reflecting the IDENTPalatal/A approach is given in (56).

(56)

		*UNSTR/o	IDENT Height	IDENT Palatal/A	*UNSTR/e	IDENTPal
(a)	*ideal* /i'de+'al/					
	☞ iðe'al				*	
	iðə'al			*!		*
	iði'al		*!			
(b)	*idealisme* /i'de+'al+'iz+m/					
	iðəə'lizmə			*!		*
	☞ iðeə'lizmə				*	
	iðəe'lizmə				*	*!*
	iðee'lizmə				**!	*
	iðiə'lizmə		*!			
(c)	*vehement* /bAA'ment/					
	☞ bəə'men					
	beə'men				*!	*
	bəe'men				*!	*
	bəi'men		*!			*
	biə'men		*!	*		*
(d)	*portem* /'pɔRt+'ɛ+m/					
	por'tɛm	*!				
	☞ pur'tɛm		*			

2.3.10 *Exceptions to eastern Catalan vowel reduction, 2: adjacent mid vowels*

In eastern Catalan the avoidance of vowel reduction of mid front vowels when the output would be *[ə'a], *['aə], *[əə] is categorical. With other sequences of mid vowels, the result is variable; avoidance of vowel reduction is predominant (Ruaix 1986; Recasens 1993; Badia i Cardús 2001), though informants suggest that variation is more finely balanced with front mid-vowels than with back mid vowels. The examples in (57) show non-centralization of an unstressed underlying mid front vowel when adjacent to a stressed mid vowel, rounded or unrounded.

(57) *deessa* [deˈɛsə] (∼ [dəˈɛsə]) 'goddess' cf. *déu* [dew] 'god'
teòleg [teˈɔlək] 'theologian' cf. *ateu* [əˈtew] 'godless'
creença [kɾeˈɛnsə] cf. *creu* [kɾɛw]
(∼ [kɾəˈɛnsə]) 'belief' 'believe.3SG.PRS.IND'
nucleó [nukleˈo] 'nucleon' cf. *nucleic* [nuˈklɛjk] 'nucleic'
conreem [kunreˈɛm] (∼ [-əˈɛm]) cf. *conreu* [kunˈrɛw] 'cultivation'
'cultivate.1PL.PRS.IND' *conrea* [kunˈrɛə]
conreés [kunreˈes] 'cultivate.3SG.PRS. IND'[31]
'cultivate.1SG.PST.SUBJ'
anorreés [ənureˈes] cf. *no re* [noˈrɛ] 'nothing'
'annihilate.1SG.PST.SUBJ'

It is in fact quite difficult to find derivational evidence such as that in (57) for underlying /e/ or /ɛ/ in examples with surface sequences [eˈo], [eˈɔ], though such surface sequences are not in themselves unusual. The fact that the proportion which is clearly derivationally related is small indicates that 'preservation' of [e] has generally to be interpreted as a matter rather of input faithfulness than of paradigm uniformity or output–output correspondence. Further examples without supporting evidence for underlying /e/ or /ɛ/ are given in (58a). Examples (58b) show the much rarer rounded + unrounded sequence.

(58) (a) *arqueòleg* [ərkeˈɔlək] 'archaeologist'
camaleó [kəməleˈo] ∼ [kəmələˈo] 'chameleon'
meteor [mətеˈɔr] ∼ [mətəˈɔr] ∼ [mətеˈor] 'meteor'
orfeó [urfeˈo] ∼ [urfəˈo] 'choral society'
peó [peˈo] ∼ [pəˈo] 'pawn'
peònia [peˈɔnjə] ∼ [pəˈɔnjə] 'peony'
reobro [reˈɔβɾu] ∼ [rəˈɔβɾu] 'reopen.1SG.PRS.IND'
reomplo [reˈomplu] ∼ [rəˈomplu] 'refill.1SG.PRS.IND'
teòric [teˈɔɾik] ∼ [təˈɔɾik] 'theoretical'
(b) *bòer* [ˈbɔer] ∼ [ˈbɔər] 'Boer'
Cloe [ˈkloe] 'Chloe'

Alongside such examples (57) and (58), with variable unstressed [e] adjacent to a stressed mid vowel, there are a good number with categorical [ə] in this environment, as in (59).

(59) (a) *aorta* [əˈɔrtə] 'aorta'
braó [bɾəˈo] 'upper arm'
braol [bɾəˈɔl] 'bellow'
caòtic [kəˈɔtik] 'chaotic'
esglaó [əzɣləˈo] 'step'
faraó [fərəˈo] 'Pharaoh'
lleó [ʎəˈo] 'lion'
Maó [məˈo] (toponym, capital of Minorca)
paó [pəˈo] 'peacock'
raó [rəˈo] 'reason'
salaó [sələˈo] 'salting'
saó [səˈo] 'maturity'

[31] Other verbs in *-ear* are similar to *conrear* 'cultivate' and *anorrear* 'annihilate'; they have some present tense forms where the /e/ or /ɛ/ in the root is stressed.

(b) *agraeix* [əɣɾəˈɛʃ] 'thank.3SG.PRS.IND'
beneeix [bənəˈɛʃ] 'bless.3SG.PRS.IND'
obeeix [uβəˈɛʃ] 'obey.3SG.PRS.IND'
paeix [pəˈɛʃ] 'digest.3SG.PRS.IND'
and similarly other verbs like these with infinitives in *-air* and *-eir*

Just as in the case in §2.3.9 of *ideal* [iðeˈal] 'ideal' vs. *Baal* [bəˈal], etc., it seems in examples (57–9) that an underlying contrast, between /A/ and /e/ or /ɛ/, rather than being neutralized as is usual in unstressed syllables, is maintained, albeit variably, adjacent to a stressed mid vowel. The contrast is generally reflected in the orthography, though *lleó* has categorical [ə] just like *raó*; and all the verbs like those in (59b) have [ə], whether spelt with *-e-* or *-a-*. So, as in §2.3.9, the faithfulness constraint IDENTPal makes its presence felt.

Generalizing from the constraint IDENTPalatal/A (55) above, one might have a contextual faithfulness constraint IDENTPalatal/$V_{[-hi]}$ (60).

(60) IDENTPalatal/$V_{[-hi]}$ (IDPAL/$V_{[-hi]}$): Adjacent to a non-High vowel, an input Palatal vowel remains Palatal in the output

The justification for IDENTPalatal/$V_{[-hi]}$ would be the difficulty of maintaining the perceptibility of [ə], the least prominent vowel, adjacent to any other vowel that is perceptually similar, i.e. one that shares with [ə] the value Mid or A. IDENTPalatal/$V_{[-hi]}$ penalizes [əe], [əɛ], [əo], [əɔ], [əə], [əa], and likewise [eə], [ɛə], [oə], [ɔə], [aə], when [ə] corresponds to a Palatal vowel in the input. (When [ə] is not related to a Palatal vowel, however, *UNSTR/e and MAX show their effects; the disfavoured sequences are better than unfaithful Palatal, or deletion.) IDENTPalatal/$V_{[-hi]}$ combines three constraints which are separable in different varieties of eastern Catalan. The first of these, IDENTPalatal/A, categorically dominates *UNSTR/e in eastern Catalan; the second, IDENTPalatal/Labial$_{[-hi]}$, dominates *UNSTR/e only for those speakers who prefer [teˈɔɾik] to [təˈɔɾik] for *teòric* 'theoretical'; and the third, IDENTPalatal/Palatal$_{[-hi]}$, dominates *UNSTR/e only for that subset who prefer [kɾeˈɛnsə] to [kɾəˈɛnsə] for *creença* 'belief'.

Tableau (61) illustrates the relevant constraint ranking for those varieties least favourable to [ə] adjacent to a non-high vowel, using examples from (57–9): (61a) *orfeó* 'choral society', (61b) *creença* 'belief', (61c) *paó* 'peacock', and (61d) *paeix* 'digest.3SG.PRS.IND'.

(61)

	orfeó	IDENT Height	IDENT Pal/V[×hi]	*UNSTR/e	IDENTPal
(a)	☞ urfeˈo			*	
	urfəˈo		*!		*
	urfiˈo	*!			
(b)	*creença*				
	☞ kɾeˈɛnsə			*	
	kɾəˈɛnsə		*!		*
	kɾiˈɛnsə	*!			
(c)	*paó*				
	peˈo			*!	*
	☞ pəˈo				
	piˈo	*!			
(d)	*paeix*				
	peˈɛʃ			*!	*
	☞ pəˈɛʃ				
	piˈɛʃ	*!			

*UNSTR/o » IDENTHeight » IDENTPalatal/V[×hi] » *UNSTR/e » IDENTPal

Sequences of underlying mid vowels both of which are unstressed are not illustrated in (61). In such cases too there is variation in eastern Catalan, for front+back sequences[32] (examples in Ruaix 1986; Bruguera 1990; Recasens 1993; Bonet & Lloret 1998; Badia i Cardús 2001). This is more surprising, since a pronunciation with regular reduction of both elements, such as *teoria* [təuˈɾiə] 'theory' (cf. *teòric* [teˈɔɾik] 'theoretical'), avoids a sequence of [ə] adjacent to a non-high vowel, via raising of unstressed /o/, /ɔ/ to [u]. In comparison, the pronunciation [teuˈɾiə] might seem a gratuitous violation of *UNSTR/e. There is no general avoidance of unstressed [əu] (e.g. *maonesa* [məuˈnɛzə] 'mayonnaise'), and stressed [ˈew] and [ˈɛw] reduce regularly to [əw] (e.g. *greu* [ˈgɾew] 'serious' – *agreujar* [əɣɾəwˈʒa] 'aggravate'; *deute* [ˈdɛwtə] 'debt' – *deutor* [dəwˈto] 'debtor'), so there can be no general constraint against centralizing a mid front vowel when followed by a labial vocoid. It is rather more likely that [e] in [teuˈɾiə], etc. is 'parasitic' on the [e] in [teˈɔɾik], etc. That is to say, in

[32] Back+front sequences have regular reduction of both vowels, giving [uə], as in: *bohemià* 'Bohemian', *coetani* 'contemporary', *proemial* 'introductory'.

this case the (optional) non-reduction is an output–output correspondence effect. And observe that I argue elsewhere (§3.2.3) that hiatus, as opposed to glide formation, in [təuˈɾiə], [teuˈɾiə], [məuˈnɛzə], and so on, is indeed an output–output correspondence effect: the labial vowel is stressed in base forms, and retains its syllabicity in derived forms. I give further examples in (62a). In (62b) are a few examples of non-reduction in words lacking obvious [eˈo] or [eˈɔ] bases, where one would have to attribute the non-reduction to a more general surface analogy.

(62)	(a)	*alveolar* [əlβeuˈlar] 'alveolar'	*alveòlisi* [əlβeˈɔlizi] 'loss of an alveolus'
		arqueologia [ərkeuluˈʒiə] 'archaeology'	*arqueòleg* [ərkeˈɔlək] 'archaeologist'
		geografia [ʒeuɣɾəˈfiə] 'geography'	*geògraf* [ʒeˈɔɣɾəf] 'geographer'
		georgià [ʒeurˈʒja] 'Georgian'	*Geòrgia* [ʒeˈɔrʒjə] 'Georgia'
		ideologia [iðeuluˈʒiə] 'ideology'	*ideòleg* [iðeˈɔlək] 'ideologist'
		meteorologia [məteuɾuluˈʒiə] 'meteorology'	*meteor* [məteˈɔr] 'meteor' – *meteoròleg* [məteuˈɾɔlək] 'meteorologist'
		teocràcia [teuˈkɾasjə] 'theocracy'	*teòcrata* [teˈɔkɾətə] 'theocrat'
	(b)	*fraseologia* [frəzeuluˈʒiə] 'phraseology'	(analogy with numerous *-eologia* words)
		neologisme [neuluˈʒizmə] 'neologism'	(analogy with *neo-* stressed initial compounding element, and/or words like *neòfit* 'neophyte')

I formulate the relevant correspondence constraint in (63). Remember that the application of IDB-DPreton[e] is variable; it is reported to be less applied in more familiar words (among which *teoria* 'theory' is usually mentioned; others are *teologia* [təuluˈʒiə] 'theology', *orfeonista* [urfəuˈnistə] 'member of a choral society').

(63) IDENT(Base, Derivative)Pretonic[e] (IDB-DPreton[e]): Pretonic [e] in a base has an identical correspondent in a derived form.

Fragment of constraint hierarchy: IDENTHeight » IDB-DPreton[e] » *OCP/V_{Mid}, *OCP/V_A, *UNSTR/e

The tableau in (64) shows the evaluation of candidates for *arqueologia* 'archaeology'. Candidates with glide formation are not considered here: see §3.2.3.

(64)

arqueologia cf. *arqueòleg*	*UNSTR/o	IDENT Height	IDB-DPreton[e]	IDENTPAL /V[xhi]	*UNSTR/e
ərkeolu'ʒiə	*!	*			*
☞ ərkeulu'ʒiə		**			*
ərk əulu'ʒiə		**	*!		
ərk ɔolu'ʒiə	*!	*	*	*	
ərkiulu'ʒiə		***!	*		

*UNSTR/o » IDENTHeight » IDB-DPreton[e], IDENTPAL/V[xhi] » *UNSTR/e

In those varieties in which IDB-DPreton[e] is inactive, [ərkəulu'ʒiə] would win over [ərkeulu'ʒiə], having no violation of *UNSTR/e.

2.3.11 *Exceptions to eastern Catalan vowel reduction 3: post-tonic* [e], [o]

Large numbers of words in eastern Catalan display regular reduction of post-tonic /e/ or /o/ (assuming these particular underlying vowels on the basis of western pronunciation, or of Majorcan pronunciation, in the case of /o/), as in the examples of (65). However, there are also a significant number of words with variable 'retention' of post-tonic [e] or [o].[33] The words involved are fairly obviously loans, from Greco-Latin, Spanish, and English sources in particular. (It may be that they are words acquired especially in educational and/or mass media contexts, where Spanish has been predominant. Non-reduction may be a means of marking their 'alien' quality, though the realization of consonants, and of unstressed *-a-*, is adapted to the regular phonological system.) In (66) are some examples (from Ruaix 1986, Bonet & Lloret 1998, and Badia i Cardús 2001) in which [e] and [o] are not supported by morphological alternation.

(65) *escàndol* [əs'kandul] 'scandal' *jove* ['ʒoβə] 'young'
febre ['feβrə] 'fever' *toro* ['tɔɾu] 'bull'

(66) Variable post-tonic non-reduced [e], [o] without stressed derivatives
àlg[e]*bra* 'algebra' *eg*[o]* 'ego'
*àloe** ['aloe] 'aloe' *estèr*[eo]* 'stereo'
apòc[o]*pe* 'apocope' *laring*[e] 'larynx'
bas[e] 'basis' *òp*[e]*ra* 'opera'

[33] Pretonic exceptions (Ruaix 1986 mentions a few, like *cereal* [seɾe'al] ~ [səɾe'al] 'cereal') are much less common, always variable, and usually attributed to language interference.

bàsqu[e]*t* ‘basketball’	*piràmid*[e] ‘pyramid’
Bost[o]*n*	*plànct*[o]*n* ‘plankton’
càt[e]*dra* ‘university chair’	*rèc*[o]*rd** ‘record’
class[e] ‘class’	*tún*[e]*l* ‘tunnel’
claudàt[o]*r** ‘square bracket’	*wàt*[e]*r* ‘WC’
cràt[e]*r* ‘crater’	*dolm*[e]*n* ‘dolmen’

*with post-tonic [e], [o] in Bruguera (1990)

Faced with just the types in (65) and (66) one might be tempted to claim that the contrast is merely one between underlying /A/ or /u/ in (65) (despite /e/ and /o/ in other dialects) vs. underlying /e/ or /o/ in (66), and that eastern vowel reduction is restricted to pretonic syllables. However, there is a sufficient number of words where derivational relation shows that post-tonic ‘reduced’ [ə] or [u] can correspond to a mid vowel (stressed) in related forms (67).

(67) Post-tonic [u], [ə] with derivational evidence for /O/ or /E/

apòstol ‘apostle’	cf. *apostòlic* ‘apostolic’
càrrega ‘load’	cf. *carrega* [kəˈrɛgə] ‘load.3SG.PRS.IND’
cèrcol ‘hoop’	cf. *cercola* [-ˈo-] ‘fit.hoops.3SG.PRS.IND’
equívoc ‘ambiguous’	cf. *equivoca* [-ˈo-] ‘confuse.3SG.PRS.IND’
fòsfor ‘phosphorus’	cf. *fosfòric* ‘phosphoric’
hidròfob ‘hydrophobic’	cf. *hidrofòbia* ‘hydrophobia’
mànec ‘handle’	cf. *emmanega* [-ˈɛ-] ‘fit.handle.3SG.PRS.IND’
micròfon ‘microphone’	cf. *microfònic* ‘microphonic’
rètol ‘label’	cf. *retola* [-ˈɔ-] ‘label.3SG.PRS.IND’
terratrèmol ‘earthquake’	cf. *tremola* [-ˈo-] ‘tremble.3SG.PRS.IND’

Just in case it might be thought that what (67) vs. (66) showed was an anti-identity base–derivative correspondence effect (with reduction when there is a derivationally related mid vowel and no reduction when there is not), we can consider examples such as those in (68), with variable retention or reduction of post-tonic [e]/[o], along with derivational relations.

(68) Post-tonic variable [e] ∼ [ə], [o] ∼ [u] with derivational evidence for /O/ or /E/

cànon ‘canon’	cf. *canònic* ‘canonical’
*caos** ‘chaos’	cf. *caòtic* ‘chaotic’
cine ‘cinema’	cf. *cinema* [siˈnɛmə] ‘cinema’
còlera ‘cholera’	cf. *colèric* ‘choleric’
*cosmos** ‘cosmos’	cf. *cosmògraf* ‘cosmographer’
*eros** ‘Eros’	cf. *eròtic* ‘erotic’
*Víctor** (given name)	cf. *Victòria*

*with post-tonic [e], [o] in Bruguera (1990)

Examples (65–8) thus show that all four squares of the truth table (69) are filled. It is possible, but not necessary, for post-tonic underlying mid vowels to undergo reduction.

(69)

	No derivational evidence for underlying mid vowel	Derivational evidence for underlying mid vowel
Reduction of post-tonic vowels	(65)	(67)
(Variable) non-reduction of post-tonic vowels	(66)	(68)

It remains for me to explain why I have insisted on limiting the context for variable non-reduction to post-tonic environments. The reason, as pointed out by Bonet & Lloret (1998: 41–2), is that when stressed suffixes are added to stems of the type seen in (66) and (68), so that the stem vowels themselves become pretonic, vowel reduction works in the regular way, as the examples in (70) show.

(70) *bostonià* [bustunˈja] 'Bostonian' cf. *Boston* [ˈboston]
catedràtic [kətəˈðɾatik] 'professor' cf. *càtedra* [ˈkateðɾə] ~ [ˈkatəðɾə] 'chair'
egoista [əɣuˈistə] 'egoist' cf. *ego* [ˈeɣo] 'ego'
opereta [upəˈɾɛtə] 'operetta' cf. *òpera* [ˈɔperə] ~ [ˈɔpəɾə] 'opera'

What, then, are the possibilities for a more formal account of variable post-tonic non-reduction of mid vowels? The radical solution is to deny that post-tonic vowel reduction is a productive synchronic process, and thus to propose that 'reduction' in examples such as (67) and (68) reflects allomorphy. Thus those who say *colèric* [kuˈlɛɾik] but *còlera* [ˈkɔlərə] would have stem allomorphs /ˈkɔlAɾ+/ and /kɔlɛɾ+/ (or /kulɛɾ+/). This interpretation seems not to appeal to native speakers, or at least not to native linguists. The number of lexical items consistent with regular post-tonic vowel reduction is very large; the lexical exceptions to it are not part of the basic vocabulary, and non-reduction is regarded at least to some degree as an interference phenomenon which can be avoided by 'careful' speakers. From this perspective, one might propose a lexically governed faithfulness constraint preserving underlying /e/ and /o/, in post-tonic syllables only, in listed words. The question of why only post-tonic vowels are affected remains open to further investigation. It is plausibly a prominence phenomenon, whereby more prominent unstressed [e] and [o] are licensed only in the relatively prominent final syllable (or foot).

2.3.12 *Further non-reduction of unstressed* [e] *in Majorca*

Majorcan displays cases of unstressed [e] in contexts beyond those mentioned in §§2.3.10–11. Detailed investigations have not been carried out, but according to Bibiloni (1998) the presence of [e] is affected by morphological relations, phonetic context, and Spanish L2 interference, with these three factors often working in combination. One context that is purely phonetic is in the post-tonic word ending [-ec] (recall that dorsal consonants are markedly palatal in much of Majorca). Bibiloni reports that [-ec] (or [-ek]) is more common than [-ək] in examples such as those in (71). In constraint terms we might say IDENTPal/_k » *UNSTR/e.

(71) *càrrec* 'load' *mànec* 'handle'
càvec 'hoe' *pròleg* 'prologue'
diàleg 'dialogue' *ràvec* 'radish'
espàrec 'asparagus'

It will be shown shortly that a following dorsal contributes to a raising environment elsewhere, too. Secondly, in less traditional words (acquired in school or mass media contexts) not only is [e] general in post-tonic syllables where there is <-*e*-> in the spelling, as in the cases of §2.3.11, but [e] occurs also in pretonic position, systematically after a labial consonant, as in *especial* 'special', *periodista* 'journalist', *aberració* 'aberration', *benzina* 'petrol', *americà* 'American', *mentalitat* 'mentality', *feliç* 'fortunate', *vehicle* 'vehicle', *vertical* 'vertical'. But 'older' words have [ə] in this environment: *pedaç* 'patch', *pessigar* 'to pinch', *besada* 'kiss' (cf. *besa* ['bezə] 'kiss.3SG.PRS.IND'), *bessó* 'twin', *mentida* 'lie', *mesura* 'measure'. (Bibiloni notes too that preceding labial protection of /e/ can spread to the next syllable, as in *f*[e]*d*[e]*ració* 'federation', *p*[e]*n*[e]*trar* 'penetrate', *v*[e]*rt*[e]*brat* 'vertebrate'.) In constraint terms, IDENTPal/Lab_(learned words) » *UNSTR/e. This constraint seems to operate variably, and to be conditioned by lexicon and register.

It is known that in certain first-conjugation verbs in Majorcan underlying root /e/ (manifest in the singular and third plural of the present indicative and subjunctive) surfaces, variably, as unreduced unstressed [e]. In his article Bibiloni (1998) mentions sixty-nine first conjugation verbs, among which thirty-five display variable unstressed [e],[34] while the remaining thirty-four have 'regular' [ə] in unstressed position. Bibiloni's account of the difference between the two verb classes goes as follows:

In principle the difference between one group of verbs and the other is an issue of chronology: the verbs with [e] are the ones with roots well established in popular usage,

[34] One of them is *f*[e]*d*[e]*rar* 'to federate', whose second, stressable [e] can best be accounted for on the basis mentioned in the previous paragraph, i.e. harmony with the preceding [e] induced by preceding labial in a 'learned' word. *Federar* is not considered further in the discussion in the text.

whereas the verbs with [ə] are of more recent introduction or diffusion. Some unresolved problems remain, however: [there follow six 'popular' verbs with [ə], four 'new' words with [e]]. (Bibiloni 1998: 534 (trans. MWW))

Bibiloni does not remark on the fact that the situation he presents of pretonic [e] in 'popular' verbs and [ə] in 'neologisms' is the reverse of that mentioned in connection with nouns and adjectives in the previous paragraph where the 'new words' like *benzina* have pretonic [e] and the old ones like *besada* have pretonic [ə]. Bibiloni also does not consider in this context the possibility that the distinction in the verbs is phonological, though he is quite willing to consider phonological influences on [e]-retention in words of other types, as we have seen. The sixty-five verbs mentioned by Bibiloni are not presented as a comprehensive survey; but if one takes them as representative, if only so as to illustrate the issues, it is possible to infer a number of ranked constraints, some of which are probabilistic. I present these constraints below in (72), along with the examples supporting them.

(72) Reduction versus (variable) non-reduction of /e/ in the final syllable of conjugation I verb roots in Majorcan

(1) Categorical reduction: before [ʎ]: *afusellar* 'shoot', *atropellar* 'hasten'

(2) Non-reduction: after palatal, after dorsal, after [n]:

after palatal	*aixecar* 'put up', *engegar* 'start', *llepar* 'lick', *llevar* 'take off', *manllevar* 'borrow'
after dorsal	*esqueixar* 'tear', *quedar* 'stay'
after [n]	*llenegar* 'slip', *nedar* 'swim', *negar* 'deny', *nevar* 'snow', *penetrar* 'penetrate'

(3) Non-reduction before palatal: *deixar* 'leave', *emprenyar* 'annoy', *engreixar* 'put on weight', *penjar* 'hang'

(4) For contexts not included under 1–3, one can estimate, on the basis of the distribution in the table below, the following contextual probabilities of non-reduction, based on coarticulatory influences of preceding or following labial, or following dorsal:

in initial syllable 1
non-initial syllable after bilabial 0.71
non-initial syllable after labiodental 0
non-initial syllable before dorsal 0.33
non-initial syllable before labial 0

	Initial non-reduction $p_e = 1$	Non-initial	
		Non-reduction	Reduction
after bilabial + before dorsal	*pegar* 'beat', *pecar* 'sin'		
after bilabial $p_e = 0.71$	*pesar* 'grieve'	*esperar* 'hope', *operar* 'operate', *imperar* 'rule', *superar* 'overcome', *recuperar* 'recuperate'	*empedrar* 'pave', *rebentar* 'burst'
after labiodental $p_e = 0$	*vedar* 'forbid', *ventar* 'fan', *vessar* 'pour'		*confessar* 'confess', *travessar* 'cross'
before dorsal $p_e = 0.33$	*pregar* 'request', *regar* 'water', *segar* 'reap'	*alegrar* 'cheer' *aregar* 'tame'	*alegar* 'allege', *arreglar* 'sort out', *delegar* 'delegate', *integrar* 'integrate'
before labial $p_e = 0$	*cremar* 'burn'		*acceptar* 'accept', *celebrar* 'celebrate', *discrepar* 'differ'

Reduction is normal in contexts other than those mentioned, that is, with adjacent dental or alveolar articulation, as in

accelerar 'accelerate'
alterar 'alter'
berenar 'have lunch'
centrar 'centre'
completar 'complete'
considerar 'consider'
contestar 'answer'
elevar 'raise'
entrenar 'train'
expressar 'express'
ingressar 'pay in'
moderar 'moderate'
molestar 'bother'
ordenar 'order'
processar 'process'
resar 'pray'
tolerar 'tolerate'

The only example of Bibiloni's unaccounted for by the principles mentioned so far is non-reduction in *fr*[e]*nar* 'brake'. Perhaps, in comparison with the eighteen examples just mentioned, one should estimate that the probability of non-reduction in the absence of favouring phonological factors is approximately 0.05. The contextual phonetic factors favouring the half-close mid front vowel in unstressed position, in particular, preceding or following palatals, and labials and dorsals among which preceding consonants are more influential than following ones, are broadly in line with what was observed in §2.2.1.3 concerning conditions on the distribution of /e/ and /ɛ/ in stressed syllables in central Catalan. In Majorcan, though, the raising/fronting effect of labials is more marked, as is

that of dorsals, the latter fact being no doubt connected with the strong tendency towards palatal realization of dorsals in Majorca. Non-reduction after /n/ is more surprising, as is reduction before /ʎ/. Rather than the presence of /ʎ/, it may in fact be the trisyllabic nature of the stem that favours [ə] in *afusellar* 'shoot' and *atropellar* 'hasten'. Summarizing and abbreviating the paradigm uniformity constraint one can say (73):

(73) (Majorcan) PUIDENTPal/ConjI/PhonContexts: A palatal vowel in the stem of an inflected conjugation I verb when unstressed corresponds to a palatal vowel in the same stem when stressed, provided certain phonological conditions involving the adjacent consonants are fulfilled.

Constraint (73) also operates variably: speakers may suppress it below *UNSTR/e. It is also in part probabilistic, though whether the probability describes the occurrence of [e] in utterances or the likelihood of a speaker assigning a word to a lexical class remains to be investigated. Bibiloni specifically mentions variation within speakers, so the former interpretation at least needs to be considered.

In nominal stems which are unstressed, underlying /e/ is retained on the whole before only the most productive and general suffixes. Such suffixes are, for example: /+ˈət/ (Diminutive), /+ˈɔt/, /+ˈas/, (both Augmentative/Pejorative), /+ˈəzə/ (de-adjectival nominal), /+ˈadə/ (action nominal). Examples: *p*[e]*dreta* 'stone.DIM', *p*[e]*ixot* 'fish.AUG', *v*[e]*llesa* 'old age', *v*[e]*ntada* 'gust of wind'. The same stems, with less productive or less general suffixes, show reduction, for example *p*[ə]*drera* 'quarry', *p*[ə]*ixater* 'fishmonger'.

Only in a yet more restricted context does underlying /ɛ/ in nominals also surface as unstressed [e] in Majorcan: the suffix must be one of the suffixes promoting paradigm uniformity *and* the preceding consonant must be bilabial. So, *Pep* [ˈpɛp] (hypocoristic of *Josep*) – *Pepet* [peˈpət] 'id.DIM'; *perla* [ˈpɛɾlə] 'pearl' – *perleta* [peɾˈlətə] 'pearl.DIM', but *gerra* [ˈʒɛrə] 'jug' – *gerreta* [ʒəˈrətə] 'jug.DIM'. The output–output correspondence constraint affecting /e/ can be formulated as follows (74), outranking *UNSTR/e.

(74) (Majorcan) IDENTB-DPal/_ProdAff: A palatal vowel in a nominal stem, unstressed, before a productive affix corresponds to a stressed palatal vowel in base.

The more narrowly constrained output–output correspondence constraint that includes /ɛ/ outranks IDENT[clo], and takes the form in (75).

(75) (Majorcan) IDENTB-DPal/Bilab_ProdAff: A palatal vowel in a nominal stem, unstressed, preceded by a bilabial, before a productive affix, corresponds to a stressed palatal vowel in the base.

Hence, the constraint hierarchy involving unstressed [e] in Majorcan, in contexts that go beyond those found in eastern Catalan in general, is as follows (76).

(76) (Majorcan) *UNSTR/ɛ, ɔ » IDENTHeight » IDENTB-DPal/Bilab_ProdAff (75) » IDENT[clo] » IDENTB-DPal/_ProdAff (74), PUIDENTPal/ConjI/PhonContexts (73) » *UNSTR/e, o » IDENTPal

In the tableau which follows, the higher-ranked constraints *UNSTR/a, *UNSTR/ɛ, ɔ » IDENTHeight, and the candidates that violate them, are omitted. Thus, only candidates showing [e] or [ə] are evaluated. The examples are *gelat* 'ice cream' /ˈʒɛl+ˈad/ from (49), in which /+ad/ is not one of the suffixes provoking B-D identity; *testet* 'pot.DIM' /ˈtest+ˈət/, where /+ˈət/ is one of the IDENTITYB-D-inducing suffixes; *perleta* 'pearl.DIM' /ˈpɛRl+ˈət+ə/; *negar* 'deny.INF (ConjI)' /ˈneg+ˈaR/ (where the verb root /ˈneg/ meets the conditions for PU (73)); *resar* 'pray.INF (ConjI)' /ˈRez+ˈaR/ (where the root /ˈRez/ does not meet the conditions for PU (73)); and *gelar* 'freeze.INF (ConjI)' /ˈʒɛl+ˈaR/. The two candidates for the last example, *gelar*, technically meet the conditions for (73) PUIDENTPal/ConjIVb/PhonContexts, (p = 1 after palatal consonant), but this is of no practical effect, since [ʒeˈla] has already been eliminated by the higher ranked IDENT[clo] inasmuch as its vowel fails to correspond to the /ɛ/ of the root.

(77)

Majorcan	IDB-D(75)	IDENT[clo]	IDB-D(74) PUID(73)	*UNSTR/e, o
gelat ʒeˈlat		*!		*
☞ ʒəˈlat				
testet ☞ tesˈtət				*
təsˈtət			*!	
perleta ☞ perˈlətə		*		*
pərˈlətə	*!			
negar ☞ neˈja				*
nəˈja			*!	
resar reˈza				*!
☞ rəˈza				
gelar ʒeˈla		*!		*
☞ ʒəˈla			*	

3

SYLLABLE STRUCTURE

Catalan syllable structure and organization has now received considerable attention from phonologists with an Optimality perspective: see, in particular, Colina (1995), Serra (1996b), Jiménez (1999), and Dols (2000). It had already become a focus of interest in post-classical generative phonology (see Montreuil 1987a; 1987b; Wheeler 1987; Harris 1993; Palmada 1993), and Lloret (2002a) provides a theoretically informed descriptive account. The presentation I offer here draws on many of the insights of this previous work, particularly the OT studies.

3.1 ONSETS

Any single consonant in the inventory may be an onset, as may a high vocoid ([j], [w]). Onset clusters consist of a consonant plus a high vocoid (Cj, Cw), or of an obstruent plus a liquid, in particular, /pɾ, bɾ, tɾ, dɾ, kɾ, gɾ, fɾ, pl, bl, kl, gl, fl/. Triconsonantal onsets, consisting of one of the obstruent+liquid clusters followed by a high vocoid, are available for some speakers, for example, in *síndria* [ˈsin.dɾjə] 'watermelon' (see §3.2.4.3). For certain more conservative speakers a *ComplexOnset constraint (*[$_{\sigma}$CCC) rules out such obstruent+liquid+glide clusters, and a high vocoid must be syllabic in this context: [ˈsin.dɾi.ə]. Many of the restrictions on onsets follow from well-known and general sonority constraints. Some other aspects, even if familiar, are not well understood. One such problem is the admissibility of /fɾ, fl/ onsets, while other strident fricatives are barred from onset clusters (*/sɾ/, */sl/, */ʃɾ/, */ʃl/), as, indeed, are the voiced counterparts */vɾ, vl/ in those varieties which have a contrastive voiced labiodental fricative.

The unmarked case in syllable organization is, of course, that all consonants may only be syllabified as margins (i.e. onsets or codas) and that all vocoids may only be syllabified as peaks (i.e. nuclei). In OT terms, the constraints excluding vowels from margins dominate all faithfulness constraints, which in turn dominate all constraints excluding consonants from margins (1). Likewise the constraints excluding consonants from peaks dominate all faithfulness constraints, which dominate all constraints excluding vowels from peaks (2) (Prince & Smolensky 1993).

(1) *M/V » Faith » *M/C

(2) *P/C » Faith » *P/V

Catalan differs from the pattern in (1) and (2) only in that, as mentioned, high vocoids may be either peaks or margins. (Their detailed distribution is the focus of §3.2.) Thus for Catalan (1) is modified so that $*M/V_{[+hi]}$ is demoted below the faithfulness constraints, as in (3).

(3) $*M/V_{[-hi]}$ » FAITH » $*M/V_{[+hi]}$ » *M/C

The syllabification constraints seen so far can be taken as an aspect of a more general unmarked syllabification algorithm, Sonority Sequence (SONSEQ), whereby in the grouping of segments into syllables sonority must increase from the beginning of an onset to the nucleus of a syllable, and must decrease from the nucleus to the end of the syllable. The outline of the Sonority Sequence (Clements 1990) is given in (4), where sonority increases from left to right on the scale.

(4) SONSEQ: sonority must increase from the beginning of an onset to the nucleus of a syllable, and must decrease from the nucleus to the end of the syllable, where the scale of sonority is:

Stops < Fricatives < Nasals < Liquids < High vocoids < Non-high vocoids

The constraint ranking of possible margins seen in (3), $*M/V_{[-hi]}$ » $*M/V_{[+hi]}$ » *M/C, formalizes a part of the Sonority Sequence constraint. Catalan grammar largely observes the Sonority Sequence constraint (except for the special case of exterior [s] in margins[1] and the morphologically conditioned case of Balearic present indicative verb forms, §8.6), and imposes some additional restrictions: not all onsets and codas that are in accord with the SONSEQ constraint are well-formed.

Grammars may impose additional requirements of minimum sonority *distance* (MSD) between adjacent segments within a syllable. Thus, for onsets, from the sequence Obstruents–Nasals–Liquids–High vocoids–Non-high vocoids Catalan allows only Obstruent+Liquid or Consonant+High vocoid sequences, while excluding Obstruent+Nasal, and Nasal+Liquid (see Colina 1995: 63 and Jiménez 1999: 67, after Clements 1990: 305). In some sense, nasals are not sufficiently different in sonority from obstruents to make an acceptable sonority gradient in an onset, nor are liquids sufficiently different in sonority from nasals. Baertsch (1998) derives sonority distance values (DIST = 1 − DIST = *n*) for onset clusters in a principled manner by Local Conjunction from Prince & Smolensky's (1993: 134–7) Peak and Margin Hierarchies. For example, an onset cluster consisting of two items from adjacent classes in the sonority hierarchy, such as /pn/ (Obstruent–Nasal), has a positive sonority distance of 1. (A Nasal–Obstruent onset sequence would have a sonority distance of −1.) By this approach, the sonority distance values depend on the number of classes that are distinguished, so, for the five classes already mentioned, Obstruents–Nasals–Liquids–High vocoids–Non-high vocoids, positive sonority distance values range from 0 to 4. In the case of Icelandic onsets, however, Baertsch proposes that finer distinctions are needed,

[1] As in *obstacle* 'hindrance' [ups.ˈtaklə], and *linx* [ˈliŋs] 'lynx'.

for example, between voiceless and voiced obstruents and between /l/ and /r/, with the result that, for this language, positive onset sonority distance values range between 0 and 8. For Catalan, adopting this approach to sonority distance directly will not give the correct result. Liquids and high vocoids are adjacent in the sonority hierarchy, and make acceptable onset clusters in Catalan, in the right order, as in *història* [is.ˈtɔ.ɾjə] 'story', *italià* [i.tə.ˈlja] 'Italian'. But onset clusters from the remaining groups which are adjacent, for example *tn, *fm, *nr, are unacceptable. While the 'sonority distance' between any consonant and a vocoid is evidently sufficient to allow a Catalan onset, among consonants only the distance between obstruents (at the least sonorous end) and liquids (at the most sonorant end) is sufficient to permit onset clusters, and not always then, as will be discussed shortly.

To summarize the situation, I provisionally propose a relevant scale of sonority with the gradient scale values as in (5). This scale and its justification is reconsidered and modified in Chapter 8, where it is shown that sonority differences drive vowel and consonant epenthesis and influence consonant gemination.

(5) Provisional Catalan sonority scale

Non-sibilant obstruents	Sibilant obstruents	Nasals	Liquids	Vocoids
1	2	3	4	7

In (5) the sonority distance between liquids, the most sonorous [+consonantal] class, and the adjacent [−consonantal] vocoids has been set at 3 (7 – 4). The class of obstruents has been subdivided into two, sibilants showing greater sonority than non-sibilants, to account for the absence of *.sl, *ʒɾ, and so on. Then the relevant Minimum Sonority Distance constraint for Catalan, MSD3, goes as follows (6):

(6) MSD3: In an onset sequence C_iC_j, the sonority value $C_j - C_i \geq 3$.

The logic of MSD is a requirement that the sonority gradient from the start of an onset to the nucleus be sufficiently steep. As formulated, MSD3 allows for triconsonantal onsets consisting of a non-sibilant obstruent, a liquid, and a glide, as in *bíblia* [ˈbi.βljə] 'bible', *pàtria* [ˈpa.tɾjə] 'motherland'. In [ˈbi.βljə], [βl] corresponds to one sequence of C_iC_j, and [lj] to a second such sequence. Of course, there is more to be said about a constraint like MSD. Why should such a constraint exist and have the characteristics it has? Some plausible suggestions can be made relating this constraint to issues of perceptibility on the one hand (Steriade 1997: 4–6; 2001b) and production difficulty on the other. Cues to consonant place and manner are generally most perceptible at a transition into a vowel (though retroflexion is most perceptible in a post-vocalic transition, as Steriade observes). For this reason CV is a 'better' syllable type than CCV as far as perceptibility of the initial consonant is concerned. From this perspective one

can see why a non-syllabic high vocoid is not significantly worse than a syllabic high vocoid as far as the perceptibility of a preceding consonant is concerned. (Of course, there may be quite other reasons why a CjV or CwV sequence might be avoided: glides are shorter than nuclei and thus have less time in which to be identified themselves.) Stop transitions into liquids share many of the characteristics of stop transitions into vowels. For example, stop VOT cues can be maintained before liquids, to a much greater degree than they can before nasals. And much the same is true of place cues in stops. On the other hand, it seems likely that the problem with sibilant+liquid onsets is not primarily a question of perceptibility—sibilants have high perceptibility in any context. Rather, such onsets demand sequencing of inconsistent gestures of the front of the tongue (and see Steriade 1994: 213 on why /pɾ/, for example, is a good onset from an articulatory point of view). A different perspective, relevant to Catalan, on the syllabification of internal sibilant+sonorant clusters derives from the fact that speakers use their knowledge of word-initial well-formedness in order to make inferences about the parsing of internal clusters (Steriade 1997: §3). In Catalan all /sC-/ initials are ill-formed, and are resolved by initial epenthesis and assignment of the two consonants to coda and onset respectively, as in *estan* [əs.ˈtan] 'stay.3PL.PRS.IND' /st+a+n/, where /s/ is in the coda of the first syllable and /t/ in the onset of the second.

I have at present no clear view about why onset clusters of non-sibilant fricative + liquid (fɾ, fl, βɾ, βl, ðɾ, ɣɾ, ɣl) are just as acceptable as clusters of obstruent stop+liquid, nor about why nasal+liquid onset clusters are disfavoured.[2] As regards /fɾ/ and /fl/, Dols (2000: 293) mentions two other respects in which /f/ in Catalan patterns with stops rather than with other fricatives. Firstly, word-final /f/ is not voiced before a following vowel, as other word-final fricatives are (*buf enorme* [ˈbuf əˈnormə] 'huge puff' vs. *bus enorme* [ˈbuz əˈnormə] 'huge diver': Recasens (1991b: 196–8)). Secondly, in Majorcan and Minorcan Catalan, word-final /f/ undergoes total assimilation, in place and manner, to a following consonant, just as stops do, but unlike other (strident) fricatives (*buf gros* [ˈbug ˈgɾɔs] 'large puff' vs. *bus gros* [ˈbuz ˈɣɾɔs] ~ [ˈbuɾ ˈɣɾɔs] 'large diver'). Note that */.vl/ and */.vɾ/ onsets are excluded even in those varieties, like Balearic, which have a /v/ phoneme. Dols (2000: 295) suggests that for Majorcan, [v] and [w] are in complementary distribution, with [v] occurring before vowels and [w] elsewhere; *[.wl] is ruled out by SonSeq (4). For other varieties with /v/, /w/ and /v/ are in contrast before vowels, with neutralization in favour of [w] before consonants; *[.wl] is ruled out by SonSeq here too. The

[2] Note that nasal+liquid clusters, even as coda+onset, are very rare in Catalan. Some which do exist are historically, and perhaps partly synchronically, compounds, e.g. *enlaire* 'aloft' (< *en l'aire* 'in the air'), *enlloc* 'anywhere' (*en lloc* 'in place'), *somriure* [sum.riwrə] 'smile' (cf. *riure* 'laugh'). *Honra* 'honour' and *Enric* (given name) illustrate the rare non-compound /n.R/ sequence. Of course */.nl/ may very well be ruled out by whatever rules out */.tl/ and */.dl/, but the absence of /.ml/ still needs an explanation.

distribution of [w] and [v] still needs to be explained, of course, but the absence of */.vl/ and */.vɾ/ is not truly a special fact about onsets.

Though the Catalan sonority scale (5) refers generally to liquids, the palatal lateral /ʎ/ does not appear in onset clusters.[3] Indeed, /ʎ/ does not appear in any kinds of intramorphemic cluster, whether onset, coda, or onset+coda. I take this to be a consequence of the universal markedness of /ʎ/ as a sound type, which means that constraints such as *Cʎ and *ʎC are ranked much higher than constraints on sequences of less marked sound types.

To account for the absence of /tl/ and /dl/ (or [ðl]) onsets, two substantive approaches have been offered (apart, that is, from attempts to deal with the issue in representational terms through some variety of 'underspecification theory', such as those of Palmada & Serra (1991) and Dols (2000: 296)). The first approach, going back to Harris (1983: 32–3) and given an OT perspective by Colina (1995: 73–7), sees the problem lying in the elements of the sequence /tl/, /dl/ ∼ [ðl], being too similar: both elements in each cluster are coronal non-continuants. (Usually only the 'underlying' */.dl/ is considered, but a 'Richness of the Base' approach should prompt us to give an account of the absence of *[.ðl] also, since, for other reasons—see §10.1—[ð] might be expected to appear in an onset after any kind of continuant.) Colina invokes a constraint *ONSET [αSL, αSL] (7).

(7) *ONSET [αSL, αSL]: Avoid having the same specifications for the Supralaryngeal node (place of articulation and manner) within the same onset.

This approach is reasonable enough if one accepts the perspective, which I also call on (§10.1.6) to account for [-l.d-] being more harmonic than [-l.ð-], according to which laterals are indeed stops ([–continuant]) with respect to adjacent consonants with shared tongue articulation ([Coronal] and [Apical]) but continuants with respect to adjacent sounds of other types. Seen thus, /.tl/ and /.dl/ would be 'partial geminates' (Kirchner 1998: 126)—adjacent consonants sharing the same place and manner (though not other features such as Lateral or Nasal). Thus, in Catalan onsets, not only is a sonority gradient (MSD) required between C_i and C_j, but also a Place/Manner 'gradient', or difference. The functional basis for such a requirement could plausibly be perceptibility of contrast, which is the basis of the second substantive approach to accounting for the absence of /tl/ and /dl/ onsets, that of Serra (1996b: 215–5; after Flemming 1995: 189–92). In /tl/ and /dl/ the obstruent stops do not have their normal tongue-front release phase which provides their 'good' perceptual clue. Lacking clear labial release, they are subject to potential confusion with /kl/, /gl/, the other stop+liquid set. The

[3] In the Catalan of Ribagorça, however, /ʎ/ does appear in onsets, but in place of the /l/ found in other varieties, e.g. *pla* [ˈpʎa] 'flat'. The realization of /.Cl/ as [Cʎ], which is also inferred to be a stage in the development of Latin /planum/ 'flat' to Italian *piano* [ˈpjano], Spanish *llano* [ˈʎano], and Portuguese *chão* [ˈʃɐ̃w], is little understood. It is likely to be related to the enhancement of contrast, or the maintenance of phonetic distance, between the reflexes of Latin [l] and [ll] when length ceased to be perceptible. See Padgett (2001) and Ní Chiosáin & Padgett (2001) on lateral contrasts in Russian and Irish from the perspective of Dispersion Theory.

interpretation of [l] as [–continuant] when adjacent to a Coronal now contributes equally to accounting for the absence of *[.ðl], inasmuch as non-strident obstruents, such as /d ∼ ð/, are realized as fricatives only between continuants (see §10.1).

In initial position, then, the possible onsets of Catalan are governed by MSD3 (6), a more specific version of SonSeq (4), together with an additional restriction expressed provisionally in *Onset [αSL, αSL] (7). In medial position too the effects of MSD3 are evident, but MSD3 on its own will not impose the correct syllable division, that is to say, the syllable division which is consistent with other phonological phenomena such as [r] ∼ [ɾ] distribution (§2.1.6), stop–fricative alternation (§10.1), coda voice neutralization (Chapter 5), and place neutralization (Chapter 6), and which is also the syllable division that corresponds to the intuition of native speakers. For example, the correct syllable division of *ària* 'aria' is [ˈa.ɾjə], rather than [ˈaɾ.jə], even though both conform to SonSeq and MSD3. The additional relevant constraint is the syllable contact law (SylCon) (Vennemann 1988: 40).[4]

(8) SylCon: The final element of a syllable is not less sonorous than the initial element of an immediately following syllable.

Notice that the version of SylCon in (8) does not penalize syllable contacts between elements of equal sonority (including geminates). One way of excluding coda–onset geminates in a language would be precisely to promote a more stringent StrongSylCon (9).

(9) StrongSylCon: The final element of a syllable is *more* sonorous than the initial element of an immediately following syllable.

(10)

	ària	MSD3	SonSeq	SylCon	*Coda
(a)	i. ☞ ˈa.ɾjə				
	ii. ˈaɾ.jə			*!	*
	iii. ˈaɾj.ə		*!	*	*
(b)	*esma*				
	i. ˈɛ.zmə	*!			
	ii. ☞ ˈɛz.mə			*	*
	iii. ˈɛzm.ə		*!	*	*

[4] Vennemann's formulation (1988: 40) is 'Contact Law: A syllable contact $A^{\$}B$ is the more preferred, the less the Consonantal Strength of the offset A and the Greater the Consonantal Strength of the onset B; more precisely, the greater the characteristic difference CS(B)–CS(A) between the Consonantal Strength of B and that of A.' By 'consonantal strength' Vennemann refers to the inverse of Clements's 'sonority'.

The strongest possible SylCon constraint would, of course, be one that ruled out internal codas altogether. Actually, although Vennemann's original formulation (see note 3 above) is not in accord with conventional OT constraint formulation, it has something in its favour, inasmuch as any candidate can be evaluated by it as 'better' or 'worse' than some alternative, provided we have a measure of 'consonantal strength' that can be applied in all cases.

Tableau (10) illustrates the role of SylCon, which ranks below MSD3. The ranking MSD3 » SonSeq is inherent, as the former is a more specific version of the latter. The examples are *ària* [ˈa.ɾjə] 'aria' and *esma* [ˈɛz.mə] 'instinct'.

Attentive readers may have observed that tableau (10) does not in fact show that SylCon is crucial, since the candidate (10a.ii) that loses on it also loses to (10a.i) on *Coda. But SylCon comes into its own in preferring a geminate affricate to a stop+fricative sequence (details in §2.1.2 and §2.1.3.2–3), as in *pitja* [ˈpid͡d͡ʒə] 'press.3sg.prs.ind'. Similarly, in words like *poble* 'town' it will prefer a geminate ([ˈpɔb.blə]) to a coda stop followed by a liquid onset (*[ˈpɔb.lə])—see §8.5. It is also the constraint by which, in the presence of inflectional /+R/ after a stem-final sonorant, [-l.dɾ-], for example, is preferred to [-l.r-], as in *valdrà* 'be worth.3sg.fut' /bal+R+ˈa/ (on which see further §8.2, §8.3). In tableau (11) candidates for *pitja* are evaluated. The constraint Maxμ rules out candidates whose stem syllable is not heavy, assuming that the input requires, however it may be represented, that [d͡ʒ] be moraic, in contrast to [ʒ].

(11)

	pitja /ˈpidʒ+a/	MSD3	SonSeq	SylCon	Max	*Coda
(a)	ˈpi.dʒə	*!			*	
(b)	ˈpid.ʒə			*!		*
(c)	ˈpi.d͡ʒə				*!	
(d)	☞ ˈpid͡.d͡ʒə					*

SylCon, Max » *Coda

3.1.1 *Onset-driven resyllabification*

Within phonological phrases, 'post-lexical resyllabification' may take place. Consider the examples in (12), where the syllabification of the individual prosodic words is presented alongside the syllabification of the same prosodic words combined into phonological phrases.

(12) (a) *cap novetat* 'no news' [ˈkap] + [nu.βə.ˈtat] → [ˈkab.nu.βə.ˈtat]
(b) *cap altre* 'no other' [ˈkap] + [ˈal.tɾə] → [ˈka.ˈpal.tɾə], *[ˈkap.ˈal.tɾə]
(c) *cap límit* 'no limit' [ˈkap] + [ˈli.mit] → [ˈkab.ˈli.mit], *[ˈka.ˈbli.mit], *[ˈka.ˈβli.mit], *[ˈka.ˈpli.mit]

(d) *mà plena* 'full hand' [ˈma] + [ˈplɛ.nə] → [ˈma.ˈplɛ.nə]
(e) *subaltern* 'secondary' [ˈsup-] + [əl.ˈtɛrn] → [ˈsu.pəl.ˈtɛrn]
(f) *subratllar* 'underline' [ˈsup-] + [rəʎ.ˈʎa]→ [ˈsub.rəʎ.ˈʎa]

Example (12b) shows a simple case of a coda consonant (here [p]) of the first word being realized as the onset of the following syllable in the phrase. In (12c), however, final [p] is not resyllabified, even though a labial stop + a liquid would make a well-formed onset cluster (as illustrated in (12d) where [pl] forms the onset of the second prosodic word). The second syllable in the (12c) phrase already has an onset ([l] in *límit*), and no additional resyllabification is called for. Examples (12e) and (12f) show the same principles applying within compounds whose first element is a bound prosodic word (here *sub-*). Observe that *subratllar* is pronounced with a voiced bilabial (coda) stop, followed by a trill, illustrating a different syllabification and allophony of /b/ and /R/ from that found, for example, in *sobrietat* [su.βɾi.ə.ˈtat] ∼ [su.βɾjə.ˈtat] 'sobriety', where /-bR-/ is intramorphemic.

Within OT phonology there are two types of approach to resyllabification of this kind, where the syllabification of phonological phrases does not exactly match that found within prosodic words (Jiménez 1999: 132–40). The first is in terms of Alignment; the alternative is a Correspondence or Paradigm Uniformity approach. The Alignment approach to the question is used, in analyses of Catalan data, by Colina (1995: 31) and Dols (2000: 298). The basic idea is that Alignment constraints, such as ALIGN(PrWd, L, σ, L) and ALIGN(PrWd, R, σ, R)[5] dominate SYLCON (8), which is relevant to word-internal syllabification. The Alignment constraints themselves, however, are dominated by ONSET.[6] This approach is illustrated in tableau (13) using examples (12b) and (12c). A prosodic word edge is indicated with '|'.

(13)

	cap altre	ONSET	ALIGNL (PrWd, σ)	ALIGNR (PrWd, σ)	SYLCON
(a)	☞ ˈka.ˈp|al.tɾə		C, VC	C, VC	
	ˈkap|.ˈal.tɾə	*!			*
(b)	*cap límit*				
	ˈka.ˈβ|li.mit		C!, CV	C, CV	
	☞ ˈkab|.ˈli.mit				*

ONSET » ALIGNL(PrWd,σ), ALIGNR(PrWd,σ) » SYLCON

[5] In fact Colina talks of aligning stems, while Dols aligns roots; but it is evident that Prosodic Word is the relevant category. A prosodic word may begin with a derivational prefix, but is treated no differently from words whose left element is a stem or root.

[6] ONSET here is an abbreviation for 'constraints of the ONSET family'.

In (13) Alignment violations are assessed in terms of consonant (C) and vowels (V) standing between the edge of the prosodic word and the nearest syllable edge both leftwards and rightwards. Thus, in the winning candidate of (13a), [ˈka.ˈpal.tɾə], the left edge of *altre* is separated by one consonant ([p]) from the left edge of the syllable it stands in, and by a vowel and a consonant ([al]) from the left edge of the next syllable. Both possibilities are considered here, given the standard interpretation of quantification in Alignment constraints ('the left edge of *every* prosodic word coincides with the left edge of *some* syllable').

The Alignment approach is satisfactory as long as the resyllabification of word-final consonants is at issue. However, as Jiménez points out, Alignment leads to the wrong results when vowel fusion (or elision) is the mechanism available to satisfy ONSET. This can be seen in the phrase *línia antiga* 'old line' [ˈli.njə] [ən.ˈti.ɣə] → [ˈli.njən.ˈti.ɣə], illustrated in (14).

(14)

	línia antiga	ONSET	ALIGNL (PrWd,σ)	ALIGNR (PrWd,σ)	SYLCON
(a)	ˈli.nj ə.ən.ˈti.ɣə	*!			
(b)	☞ ˈli.njən.ˈti.ɣə		C!C, VC	CCV, C	
(c)	☜ ˈlin.jən.ˈti.ɣə		C, VC	CV, C	*

In (14) I have assumed that the two adjacent cases of [ə] in *línia antiga* [ˈli.njə.ən.ˈti.ɣə] are fused into one to comply with ONSET. (Jiménez attributes the fusion to OCP violation, but this seems unnecessary, since adjacent unstressed vowels at word edges do not have to be identical, as in this particular example, to provoke loss, or glide formation. I assume ONSET is driving the process here as in (13).) The fact that fusion is the outcome, rather than realization as a long vowel, reflects the undominated position in Catalan of *Vμμ (No long vowels), or, more particularly, the ranking *Vμμ » MAXμ. In the 'fusion' interpretation, the [ə] in the second syllable of [ˈli.njən.ˈti.ɣə] corresponds to both the final vowel of *línia* and the initial vowel of *antiga*. Or, to put it another way, the realization of the two prosodic words overlaps.

The problem in (14) is that the Alignment constraints necessarily penalize (incorrectly in (14b)) any syllable with more elements in it than there are in some conceivable alternative (e.g. (14c)), since the more complex the onset or coda, the further some element in the syllable must be from one edge or the other. As Jiménez argues, candidates like (14b) ought rather to be preferred because, while respecting ONSET, they are maximally faithful to the syllabification of the input, which in this case consists of the individual prosodic words. Rather than by Alignment constraints, then, resyllabification is limited by what Jiménez (1999: 134) calls a SYLLABIC UNIFORMITY (UNIFORMσ) constraint, which I give here with a new formulation (15).

(15) SYLLABIC UNIFORMITY (UNIFORMσ): The syllabic organization of elements of a phonological phrase is the same as the syllabic organization of the corresponding elements in prosodic words.[7]

SYLLABIC UNIFORMITY is an output–output correspondence constraint favouring the uniform realization of words, specifically penalizing resyllabification when words are combined in utterances. Tableau (16) now replaces (13) and (14), with UNIFORMσ standing in place of ALIGN(PrWd, L, σ, L), ALIGN(PrWd, R, σ, R). MAXV is included, just to illustrate that it must rank below ONSET.

(16)

	cap altre	ONSET	UNIFORMσ	MAXV	SYLCON
(a)	☞ ˈka.ˈp|al.tɾə		*		
	ˈkap|.ˈal.tɾə	*!			*
(b)	*cap límit*				
	ˈka.ˈβ|li.mit		*!		
	☞ ˈkab|.ˈli.mit				*
(c)	*línia antiga*				
	ˈli.njɚ|.ən.ˈti.ɣə	*!			
	☞ ˈli.njən.ˈti.ɣə			*	
	ˈlin.jən.ˈti.ɣə		*!	*	*

ONSET » UNIFORMσ, MAXV » SYLCON

As before, in (16) the interpretation is that the vowel of the second syllable of [ˈli.njən.ˈti.ɣə] corresponds simultaneously to the final vowel of *línia* and to the initial vowel of *antiga*. Thus [ˈli.njən.ˈti.ɣə] maintains perfect syllabic uniformity, even though the syllable structures of the component words overlap: $[_i$ˈli.nj$[_j$ə$]_i$n.ˈti.ɣə$]_j$. It may be necessary to spell out that SYLLABIC UNIFORMITY ignores syllable boundaries at the edge of domains; otherwise, syllabification of $[_i$ˈli.nj$[_j$ə$]_i$n.ˈti.ɣə$]_j$ is not perfectly uniform. Be that as it may, the winning candidate is more 'uniform' than the loser that also respects ONSET.

[7] Jiménez formulates the constraint in terms of 'lexical elements' (morphemes, roots, affixes, words), but this implies that syllabification is a lexical ('input') property, rather than a property of phonological words or clitics—an output property. I take UNIFORMσ to be an output–output Correspondence constraint.

3.2 THE SYLLABIFICATION OF VOCOIDS

3.2.1 *Introduction*

The major focus of § 3.2 is the syllabification of high vocoids within prosodic words. In Chapter 4 I extend the analysis to sandhi environments, where the syllabification of [ə] is also investigated. The framework of the account starts with a presentation of a typologically neutral or unmarked syllabification of vocoids. I then show in what respects, and according to what principles, the facts of Catalan pronunciation deviate from this.

3.2.2 *Default syllabification of high vocoids within words*

The basic principles of syllable-structure constraints in Catalan have been presented in §3.1. A constraint ranking mentioned there is repeated here as (17). The effect of (17) is to exclude non-high vocoids from margins, while allowing high vocoids, and of course consonants, as margins.

(17) $*M/V_{[-hi]}$ » FAITH » $*M/V_{[+hi]}$ » *M/C

The constraint ranking in (17), combined with ONSET, can force high vocoids into margin positions in Catalan. This can be illustrated (18a, b) with the examples *teu* 'your.M' and *iaia* 'granny'. The undominated constraints banning consonants from peak positions are ignored here.

(18)

	teu	*M/a,ə	*M/e,o	ONSET	*M/i,u	*CODA
(a)	te.u			*!		
	☞ tew				*	*
	tẹu		*!			
(b)	*iaia*					
	i.a.i.ə			*!***		
	i.a̯iə̯	*!*		**		*
	jajə̯	*!			**	**
	jaj.ə			*!	**	*
	☞ ja.jə				**	

I also take as an unmarked starting assumption the (undominated) high ranking of a constraint excluding complex nuclei, *COMPLEXN. This constraint rules out both nuclei with more than one place of articulation (an interpretation often

offered for Spanish rising diphthongs, e.g. by Colina 1995), and nuclei with more than one mora.

In conformity with SonSeq (§3.1 (4)), adjacent consonants of equal sonority must be divided between syllables; thus *acte* [ak.tə] 'act', *himne* [im.nə] 'hymn'. A related issue that will concern us here (§3.2.5) is the syllabification of adjacent high vocoids /ii/, /uu/, /iu/, /ui/; in the absence of sonority fall or rise, their syllabification is not determined by the constraints given so far (Rosenthall 1994: 137). The scene can partially be set by mentioning the constraint disfavouring complex onsets *ComplexO,[8] which in post-consonantal position will favour [iw] over [ju] and [uj] over [wi]. The effects of the constraint ranking can be seen in (19) with the examples *buida* [ˈbujðə] 'empty.F', *riure* [ˈriwɾə] 'laugh.INF', and *iuca* [ˈjukə] 'yucca'.

(19)

		*ComplexO	Onset	*Coda
(a)	*buida*			
	bu.i.ð̞ə		*!	
	bwi.ð̞ə	*!		
	☞ buj.ð̞ə			*
(b)	*riure*			
	ri.u.ɾə		*!	
	rju.ɾə	*!		
	☞ riw.ɾə			*
(c)	*iuca*			
	i.u.kə		*!*	
	☞ ju.kə			
	iw.kə		*!	*

*ComplexO, Onset » *Coda

The final element of unmarked organization to be mentioned here is that lexically stressed elements must be syllable nuclei (a version of 'faithfulness to prosodic heads' (FPH): Beckman 1998; Alderete 1999; Kager 2000). As will be seen (Chapter 9), word stress in Catalan is contrastive, though subject to some redundancy, as might be expected. This means that a high vocoid, provided it is stressed, may be in hiatus with an adjacent non-high vocoid. This is illustrated in (20) with *veí* [bə.ˈi] 'neighbour', *tia* [ˈti.ə] 'aunt', contrasted with *Tià* [ti.ˈa] 'Seb' (hypocoristic of *Sebastià* 'Sebastian'). The evaluation of candidates for *tia, veí* shows FPH » Onset.

[8] I return later to elaborate the details of this constraint in §3.2.4.

(20)

		FPH	*ComplexO	Onset	*M/V[+hi]
(a)	*veí*				
	ˈbəj	*!			*
	☞ˈbə.ˈi			*	
(b)	*tia*				
	ˈtja	*!	*		*
	ti.ˈa	*!		*	
	☞ ˈti.ə			*	
(c)	*Tià*				
	ˈti.ə	*!		*	
	ˈtja		*		*!
	☞ ti.ˈa			*	

FPH » Onset

In §3.2.3 the syllabification of high vocoids following other vocoids is considered; we have seen two such examples, *teu* 'your.M' in (18a), and *veí* 'neighbour' in (20a). In section §3.2.4 it is the turn of high vocoids preceding other vocoids, of which examples have already been seen in *iaia* 'granny' (18b), *tia* [ˈti.ə] 'aunt' (20b), and *Tià* [ti.ˈa] 'Seb' (20c). In §3.2.5 I investigate further adjacent high vocoids not already covered in §3.2.3 and §3.2.4, i.e. high+high. (Some high+high examples have already been seen: *buida* 'empty.F', *riure* 'laugh.INF', and *iuca* 'yucca' in (19).)

3.2.3 *High vocoids following other vocoids*

The unmarked syllabification of high vocoids after non-high vocoids is that they become codas in falling diphthongs. The following examples illustrate this default syllabification in Catalan:

teu [ˈtew] 'your.M' (18a)	*flairar* [fləj.ˈɾa] 'sniff.INF'
mai [ˈmaj] 'ever'	*Europa* [əw.ˈɾɔ.pə] 'Europe'
nou [ˈnɔw] 'nine'	*compràveu* [kum.ˈpɾa.βəw]
llei [ˈʎej] 'law'	'buy.2PL.PST.IMPF'.

After an onset, a falling diphthong is also the default syllabification of a sequence of two high vocoids, as seen at (19) above; further examples (unstressed) are *cuidava* [kuj.ˈða.βə] 'care.3SG.PST.IMPF', *anéssiu* [ə.ˈne.siw] 'go.2PL.PST.SUBJ'. In this section I investigate the circumstances in which this unmarked syllabification is overridden. It has already been observed, in the example *veí* [bə.ˈi] 'neighbour' (20a), that lexically stressed high vowels are protected from glide formation when preceded by another vowel. Thus word-internal violations of

ONSET are by no means rare in Catalan. Here are a few more examples: *peülla* [pə.ˈu.ʎə] 'hoof', *diluir* [di.lu.ˈi] 'dilute.INF', *geniüt* [ʒə.ni.ˈut] 'irascible'. (Less formal or less conservative pronunciations for the last two are [di.ˈlwi], [ʒə.ˈnjut]; see §3.2.4.7.)

3.2.3.1 *Stressed vowel + high vowel hiatus*

I consider first high vowels in hiatus after stressed vowels. Observe surface contrast between falling diphthongs in (21) and hiatus in (22).

(21) Falling diphthong [ˈVj] [ˈVw]
noi [ˈnɔj] 'boy' *creu* [ˈkɾɛw] 'believe.3SG.PRS.IND'
reis [ˈrejs] 'kings' *interviu* [intəɾˈβiw] 'interview'

(22) Hiatus [ˈV.i] [ˈV.u]
lloï [ˈʎɔ.i] 'praise.1/3SG.PRS.SUBJ' *creo* [ˈkɾe.u] 'create.1SG.PRS.IND'
creïs [ˈkɾe.is] 'create.2SG.PRS.SUBJ' *canvio* [kəm.ˈbi.u] 'change.1SG.PRS.IND'

What is characteristic of the hiatus pattern in (22) is the presence of a verbal inflectional morpheme /+i/ 'Present Subjunctive' or /+u/ '1SG.Present Indicative'. These morphemes are consistently resistant to glide formation when preceded by vowel-final verb stems. The motivation appears to be to preserve the morphological integrity of these inflectional elements, or their 'paradigm uniformity' given that they are realized as vocalic nuclei in the great majority of their occurrences, i.e. after consonant-final verb stems. A way to represent this fact is through a Metrical Consistency constraint (see Burzio 1994: 314 ff. on Metrical Consistency in general).[9] In fact, SYLLABIC UNIFORMITY, mentioned above (15), is itself a type of metrical consistency constraint, one that requires that words retain their independent prosodic organization when linked in phrases. I call the relevant constraint here INFLECTIONAL CONSISTENCY (syllable structure) (23).

(23) INFLECTIONAL CONSISTENCY (syllable structure) (INFLCONSσ). The syllabic organization of an inflectional suffix is the same in all the words it appears in.

It is probably unnecessary for INFLCONSσ itself to specify what the required syllable organization is. If the suffix contains a non-high vowel, that can only be a syllabic nucleus (though INFLCONSσ might protect it from deletion). In Catalan, only high vowels can be syllabified differently in principle, as either peaks or margins. Taking /+i/ 'subjunctive' as an example: after a consonantal stem such as /pɔz/ 'to put', /i/ can only be a syllabic nucleus, as *[ˈpɔzj] falls foul of undominated SONSEQ (4). After a vowel-final stem, such as /ʎɔ/ 'to praise', the suffix /+i/

[9] I am very grateful to Maria-Rosa Lloret for suggesting to me the potential of a Metrical Consistency approach to some of the phenomena in this chapter.

could be syllabified as [j] (*[ˈʎɔj]) or as [i] ([ˈʎɔ.i]) without violating SonSeq. The correct form [ˈʎɔ.i] achieves consistent syllable-structure realization of the suffix, while displaying an Onset violation, thus demonstrating the ranking in (24).

(24) InflConsσ » Onset

InflConsσ might, however, need to be restricted to the vocoid element of a suffix. The past subjunctive affix /+es/ ends in a consonant which is sometimes a coda (as in *posés* 'put.1/3sg.pst.subj') and sometimes an onset (as in *posessin* 'put.3pl.pst.subj'). The InflConsσ constraint would presumably impose the wrong syllabification on *posessin*, namely [pu.ˈzes.in], since *posés* syllabified as [pu.ˈze.s] would violate high ranking *P/C (2), assuming that Gen requires that any syllable have a peak. I leave aside for the moment the possible further precision of InflConsσ.

Notice that vocalic verb affixes may indeed be forced into margins by higher-ranking constraints, in phrasal contexts. Thus *estudio animals* 'I study animals', consisting of [əs.tu.ˈði.u] + [ə.ni.ˈmals], may be syllabified [əs.tu.ˈði.wə.ni.ˈmals], by rhythmic constraints (of the *Lapse family; see (33) below) dominating InflConsσ.

3.2.3.2 *Hiatus before 'underlying' stress*

A second type of exception to the unmarked syllabification of falling sonority vocoid sequences ˈVj, ˈVw is illustrated in (25).

(25)	(a)	*caos* [ˈka.us] ∼ [ˈka.os] 'chaos'	*caòtic* [kə.ˈɔ.tik] 'chaotic'
		període [pə.ˈɾi.u.ðə] 'period'	*periòdic* [pə.ɾi.ˈɔ.ðik] ∼ [pə.ˈɾjɔ.ðik] 'periodical'
		etíop [ə.ˈti.up] 'Ethiopian'	*etiòpic* [ə.ti.ˈɔ.pik] ∼ [ə.ˈtjɔ.pik] 'Ethiopian'
		alvèol [əl.ˈβɛ.ul] 'alveolus'	*alveòlisi* [əl.βe.ˈɔ.li.zi] 'alveolysis'
		pecíol [pə.ˈsi.ul] 'petiole'	*peciòlul* [pə.si.ˈɔ.lul] ∼ [pə.ˈsjɔ.lul] 'petiolule'
	(b)	*nuclèol* [nu.ˈklɛ.ul] 'nucleolus'	
		trio [ˈtɾi.u] 'trio'	
		Antíoc [ən.ˈti.uk] 'Antioch'	

The alternations exemplified in (25a) make it clear that the vowel which appears as syllabic [u] after a stressed vowel in the items in the left-hand column is 'underlyingly' (to use derivational terminology) non-high. When unstressed it is realized as [+high] by the general eastern continental Catalan vowel reduction process (§2.3), but it is nonetheless 'protected' from glide formation. There is in fact no pattern where an unstressed syllabic vocoid [u] in a word of the form *etíop* [ə.ˈti.up], etc. alternates with a stressed [ˈu] in a derivationally related form, such as *[ə.ti.ˈu.pik]. As the orthography suggests, of course, the words in (25) have post-tonic [o] in Catalan varieties that reduce only unstressed low rounded vowels to mid. The fact that this surfaces as low mid [ɔ] in the

suffixed forms illustrates an independent process of neutralization favouring low mid over high mid vowels in words of this prosodic and morphological make-up (Mascaró's law: see §2.2.1.2). Stressed/unstressed vowel alternations as illustrated in (25a) provided good arguments in standard GP for relatively abstract underlying forms of morphemes, in circumstances where no one surface morphological context is able to provide all the information to determine the vowel qualities in the various derived forms; thus *período* would have the UR /paˈɾiode/, and *periòdic* would be /paˈɾiode+ik/, or something along those lines. In OT, to deal with such phenomena, a Correspondence Theory (or Metrical Consistency) approach would generally be preferred. From this perspective, the syllabic nature of [u] in *período* [pəˈɾi.uðə] would be guaranteed by a requirement to match the syllable structures of base and derivative, though in cases such as (25a) observe that the qualities of the base are required to match those of the derivative, and not the other way round. This fact may specifically favour a Metrical Consistency account over a Faithfulness to Prosodic Heads (FPH) approach which would require a prosodic head in a base to be matched by a prosodic head in its derived form. In this case Base consistency of syllable structure is at issue. The relevant output–output correspondence constraint, however expressed, would dominate the syllabification constraints such as I have proposed in (19), which would otherwise declare *[pə.ˈɾiw.ðə] the winning candidate, parallel to *viuda* [ˈbiw.ðə] 'widow'. The correspondence constraint will, in turn, be dominated by the constraints imposing vowel reduction in unstressed syllables. That is, in Catalan Base–Derivative correspondence constraints may constrain syllable organization—there is good evidence that they do—but they cannot impose matching vowel quality in defiance of the vowel reduction constraints. A correspondence account still needs to explain why, for example, in the case of a derivational pair like *caos* [ˈka.us] 'chaos', *caòtic* [kəˈɔtik] 'chaotic', the output *[ˈkaws] – *[ˈkawtik] is not more harmonic; compare *nau* [ˈnaw] 'ship' – *nàutic* [ˈnawtik] 'nautical'. In fact, the contrast in surface pattern must be lexical. To simplify, the representation /kaos/ for *caos*, in contrast with /nau/ or /naw/[10] for *nau*, abbreviates whatever this lexical contrast is. That is, /o/ stands here for 'a rounded vowel that may be stressed and non-high, and may not be a coda'.

The examples in (25b) are more problematic for a correspondence-theory account, since these words do not participate in morphological alternations (or none that suffice to account for syllabic [u]). For words like *nuclèol* [nuˈklɛ.ul] 'nucleolus', and there are a few more of this general shape, it would be plausible to claim that a suffix /+ol/ is present, the very one that shows up in *alvèol* [əlˈβɛ.ul] 'alveolus' and *pecíol* [pəˈsi.ul] 'petiole', for example. Thus the correspondence would be established between the occurrences in various contexts of

[10] According to the theory of Richness of the Base (Smolensky 1996) it is a matter of no consequence whether /nau/ or /naw/ is chosen as the input. These forms do not contrast, and the syllabification constraints as presented in this chapter allow only [naw] to surface.

this morpheme, sufficient to impose its syllabicity wherever it appeared. (And this is exactly the argument that would have led, in GP, to proposing an underlying form /+ol/, or /+Ol/, given that the underlying value of the feature [low] is indeterminate.) Such a case is harder to make for *trio* [ˈtɾi.u] 'trio', unless one argues that it contains a morpheme /+o/ found also in *duo* [ˈdu.o] 'duet'—here the adjacent [u] prevents vowel reduction of unstressed [o] to [u]. A parallel case cannot, in any case, be made for *Antíoc* [ənˈti.uk] 'Antioch'. We are obliged in such cases to regard the post-tonic vocoid as lexically marked for syllabicity. It is encouraging, though, for the present approach that such examples are vanishingly few.

The reader will have noticed that all the examples in (25) involve a rounded post-tonic vocoid. In (26) are seen some cases comparable to those of (25a), in which a stressed vowel is followed by a high front vocoid that appears in a different shape in a derivative.

(26)	*laic* [ˈlajk] 'laic'	*laïcitzar* [la.isidˈd͡za] 'laicize.INF'
	heroic [əˈɾɔjk] 'heroic'	*heroïcitat* [əɾu.isiˈtat] 'heroicness'
	fluid [ˈflujt] 'fluid'	*fluïdesa* [flu.iˈðɛzə] 'fluidity', *fluídic* [flu.ˈiðik] 'fluidic'
	druida [ˈdɾujðə] 'druid'	*druïdisme* [dɾu.iˈðizmə] 'druidism', *druídic* [dɾu.ˈiðik] 'druidic'
	tifoide [tiˈfɔjðə] 'typhoid'	*tifoïdal* [tifu.iˈðal] 'typhoid (adj.)'

In contrast with what was seen in (25), the high front vocoid in a post-tonic context (left-hand column) is never syllabic, even when there is evidence, as here in the related forms in the right-hand column, to suggest that the high front vocoid is syllabic 'underlyingly'. There are, indeed, no words in Catalan containing the string -ˈV.i, beyond the type already considered at §3.2.3.1 (22), where the present subjunctive inflectional morpheme /+i/ is involved. Words of the type illustrated in (26) contain the Greco-Latin derivational suffixes *-ic* /+ik/, *-id* /+id/ and *-oide* /+ˈɔide/. On general grounds one might hope to say that the high vocoid of these suffixes was lexically unspecified for syllabicity, so that it would turn up as syllabic after a consonant (as it in fact does, in *vàlid* 'valid', *físic* 'physical', etc.) and as a glide after a vowel (as in the left-hand column of (26) and generally in the affix *-oide*). However, the forms in the right hand column of (26) argue against this simple solution. (In fact, the affix *-oïdal* in words like *tifoïdal* 'typhoid (adj.)' is quite often pronounced [-ujˈðal], assimilating them to the default syllabification pattern.) Rather than 'underlying' syllabicity, then, one could claim that the syllabic realization of the suffixes *-ic* /+ik/ and *-id* /+id/ which is found in the vast majority of their occurrences (Recasens 1993: 109) imposes itself, due to a constraint of metrical consistency in the syllabification of a *derivational* suffix DERIVCONSσ, in the handful of cases where these suffixes are immediately preceded by an unstressed vowel (*laïcitzar* [la.isidˈd͡za], *fluïdesa*

[flu.iˈðɛzə], etc.). However, DERIVCONSσ is overridden in the post-tonic position. Then, one must ask, why this difference between post-tonic and post-atonic? And why the difference between the unrounded (26) and the rounded (25) post-tonic outcome?

The difference between post-tonic and post-atonic vocoids that shows up in (26) reflects, I argue, the unmarked state of affairs that a $\text{ˈV.V}_{[+hi]}$ hiatus (in which the first vowel is stressed) is even less harmonic than a $\text{V.V}_{[+hi]}$ hiatus in which neither element is stressed. This is because, perceptually, it is very hard to distinguish between a fall in prominence from stressed to unstressed (in adjacent syllables with hiatus $\text{ˈV.V}_{[+hi]}$) and a fall in prominence from peak to coda (within a falling diphthong $\text{ˈVV}_{[+hi]}$). In a $\text{V.V}_{[+hi]}$ hiatus where neither element is stressed it is somewhat easier to maintain a balance of prominence, and perhaps especially so when the vocoids are of equal sonority, such as they are in *heroïcitat* [ərु.isiˈtat] or *fluïdesa* [flu.iˈðɛzə]. In OT terms, there is an unmarked (Paninian) constraint ranking $\text{*ˈV.V}_{[+hi]}$ » $\text{*V.V}_{[+hi]}$ (between which other constraints may appear).

The difference between the outcome of unrounded post-tonic vocoids (26) and rounded ones (25) plausibly reflects, in correspondence constraints, the ranking of potential margin constraints that were considered at (18) above, namely *M/e,o » *M/i,u. That is to say, lexical non-high vowels are more resistant to glide formation than lexical high vowels are, other things being equal (cf. Serra 1996b: 85). In (27a) I give a relatively informal statement of the correspondence constraint imposing syllabicity matching in mid vocoids, in (27b) a more general constraint penalizing prosodic allomorphy in derivational suffixes, and in (28) the constraint against post-tonic hiatus with falling sonority.

(27) (a) MID VOWEL CONSISTENCY: (MIDCONSσ): A mid vowel in a base or derivative and its correspondent match in syllable organization. (Note that the constraint merely requires that there be a mid vowel in one of the correspondents, rather than that there be a mid vowel in the base.)

(b) DERIVATIONAL CONSISTENCY (syllable structure) (DERIVCONSσ): A derivational suffix has the same syllable structure in all the words it appears in.

(28) $\text{*ˈV.V}_{[+hi]}$: A stressed vowel is not immediately followed by a high vowel nucleus.

The constraint ranking so far is illustrated in the tableaux in (29) with (a) *caus* [ˈkaws] 'fall.2SG.PRS.IND.', (b) *caos* [ˈka.us] 'chaos', (c) *fluid* [ˈflujt] 'fluid', and (d) *fluïdesa* [flu.i.ˈðɛ.zə] 'fluidity'. The Metrical Consistency constraints outrank elements of the ONSET constraint family that disfavour hiatus.

(29)

		MIDCONSσ	*ˈV.V[+hi]	DERIVCONSσ	ONSET
(a)	*caus* / ˈkau+s/				
	ˈka.us		*!		*
	☞ ˈkaws				
(b)	*caos* /ˈkaos/ cf. *caòtic*				
	☞ ˈka.us		*		*
	ˈkaws	*!			
(c)	*fluid* /ˈflu+id/ cf. *fluídic*				
	ˈflu.it		*!		*
	☞ flujt			*	
(d)	*fluïdesa* /flu+id+ˈɛz+a/				
	☞ flu.iˈðɛzə				*
	flujˈðɛzə			*!	

MIDCONSσ » *ˈV.V[+hi] » DERIVCONSσ » ONSET

3.2.3.3 *Unstressed vowel followed by unstressed high vocoid*

I consider now sequences of vowel+high vocoid when neither element receives lexical stress. As mentioned in §3.2.2, the unmarked syllabification as a falling diphthong is well attested in Catalan. At (30) are a few more examples, chosen to illustrate atonic falling diphthongs that do not alternate with stressed diphthongs in different morphological contexts.

(30) *autor* [əw.ˈto] ‘author’ *Eivissa* [əj.ˈβi.sə] ‘Ibiza’
feudal [fəw.ˈðal] ‘feudal’ *oidà* [oj.ˈða] ‘bravo!’
ciutat [siw.ˈtat] ‘city’ *buirac* [buj.ˈɾak] ‘quiver’
faisà [fəj.ˈza] ‘pheasant’ *tròlei* [ˈtɾɔ.ləj] ‘trolley bus’
apaivagar [ə.pəj.βə.ˈɣa] ‘soothe.INF’

Hiatus is found (possibly variably) when the first vowel of the sequence ends a prefix, as in the examples in (31).

(31) *prohibir* [pɾu.i.ˈβi] ‘prohibit.INF’
reincidir [rə.in.si.ˈði] ‘relapse.INF’
coincidir [ku.in.si.ˈði] ‘coincide.INF’
prehistòric [pɾə.is.ˈtɔ.ɾik] ‘prehistoric’
preordinar [pɾə.uɾ.ði.ˈna] ‘preordain.INF’
reorganitzar [rə.uɾ.ɣə.ni͡d.ˈd͡za] ‘re-organize.INF’

These examples involving a vowel-final prefix and a vowel-initial stem also display a Metrical Consistency effect, this time relativized to stem-initial pos-

ition, in a correspondence version of positional faithfulness (to prosodic heads: FPH, cf. (20)). It can be argued that hiatus preserves the shape of the salient initial element of the stem, a strong position because of its importance to morpheme identification. I formulate the constraint in (32).

(32) INITIAL CONSISTENCY (INITCONS): The initial element of a stem which is the head of a syllable has a correspondent which is the head of a syllable in a derivative.

An alternative to INITCONS would be a morphological alignment constraint ALIGNRPrefix: ALIGN(Prefix, R, σ, R), but alignment constraints have a rather stipulative character. Morphological uniformity or consistency constraints, or output–output correspondence constraints in general, are more strongly founded on anti-allomorphy motivation.

As Recasens (1993: 110) points out, INITCONS may not be observed if the morphological structure has lost salience for speakers, as in quite common words such as *reunir* 'unite' or *preocupar* 'concern' (*reunir* is much more frequent than *unir* 'unite' and *preocupar* is not semantically transparent). A rhythmic constraint *LAPSE, disfavouring a succession of more than two unstressed syllables, may also prefer e.g. for *preordinar* 'pre-ordain.INF' [pɾəwɾ.ði.ˈna] to [pɾə.uɾ.ði.ˈna], and likewise for all but the first example in (31) where for *prohibir* 'prohibit.INF' [pɾu.i.ˈβi] is more generally preferred to [pɾuj.ˈβi].

I give a version of a suitable member of the *LAPSE family in (33), though I return to *LAPSE in more detail later (§3.2.4.6 and Chapter 9).

(33) *LAPSE3σ: A sequence of three unstressed syllables is not permitted.

Provisionally, I propose a constraint ranking as illustrated in tableau (34) with (a) *reiteradament* [rəj.tə.ˈɾa.ðə.ˈmen] 'repeatedly' (/Re+iteRa-/) and (b) *prohibir* [pɾu.i.ˈβi] 'prohibit.INF' (/pRo+ib-/).

(34)

	reiteradament	*LAPSE3σ	INITCONS	ONSET
(a)	rə.i.tə.ˈɾa.ðə.ˈmen	*!		*
	☞ rəj.tə.ˈɾa.ðə.ˈmen		*	
(b)	*prohibir*			
	☞ pɾu.i.ˈβi			*
	pɾuj.ˈβi		*!	

*LAPSE3σ » INITCONS » ONSET

Hiatus is also found in pretonic position in stems where the high vowel corresponds to a vowel that receives stress in a related form (35) (Recasens 1993: 109; Cabré & Prieto 2004.). The relation may be inflectional or derivational, but is more straightforwardly a relation of derivative to base than the cases considered at (27).

(35) (a) *saücar* [sə.u.ˈka] 'elder grove' — *saüc* [sə.ˈuk] 'elder'
veïnat [bə.i.ˈnat] 'neighbourhood' — *veí* [bə.ˈi] 'neighbour'
ruïnós [ru.i.ˈnos] 'ruinous' — *ruïna* [ru.ˈi.nə] 'ruin'
violí [bi.u.ˈli] 'violin' — *viola* [bi.ˈɔ.lə] 'viola'
geogràfic [ʒe.u.ˈɣɾa.fik] 'geographical' — *geògraf* [ʒe.ˈɔ.ɣɾəf] 'geographer'

(b) *proveïdor* [pɾu.βə.i.ˈðo] 'supplier' — *proveir* [pɾu.βə.ˈi] 'supply.INF'
enraonar [ən.rə.u.ˈna] 'speak.INF' — *enraona* [ən.rə.ˈo.nə] 'speak.3SG.PRS.IND'
vehicular [bə.i.ku.ˈlar] 'vehicular' — *vehicle* [bə.ˈi.klə] 'vehicle'
amoïnar [ə.mu.i.ˈna] 'bother.INF' — *amoïna* [ə.mu.ˈi.nə] 'bother.3SG.PRS.IND'
obeirítem [u.βə.i.ˈɾi.əm] 'obey.1PL.COND' — *obeïm* [u.βə.ˈim] 'obey.1PL.PRS.IND'
arqueològic [ər.ke.u.ˈlɔ.ʒik] 'archaeological' — *arqueòleg* [ər.ke.ˈɔ.lək] 'archaeologist'
meteorològic [mə.te.u.ɾu.ˈlɔ.ʒik] (also -te.o-, -tə.u-, -təw-) 'meteorological' — *meteor* [mə.te.ˈor] 'meteor'

The pattern in (35a, b) illustrates the operation of a Metrical Consistency correspondence constraint that requires the head of a syllable (a nucleus) in the base to correspond to the head of a syllable in the derivative (MetConsσ)—compare Faithfulness to Prosodic Heads (FPH) (20). MetConsσ (36) outranks Onset.

(36) Metrical Consistency (nucleus) (MetConsσ): The head of a syllable in a base corresponds to the head of a syllable in a derivative. (MetConsσ may well subsume InitCons.)

In (35b) are some longer words where hiatus is likewise found in more formal styles at least, reflecting MetConsσ » Onset. In several of the examples of (35) displaying [u] – [ˈɔ] alternations, what is shown illustrates another effect of part of the constraint ranking presented at (29) above: MidConsσ » Onset. The Metrical Consistency constraint enforces syllabicity of the 'underlying' mid vowel /ɔ/ over the output of the competing syllabification constraint that favours diphthongs over hiatuses.

As was the case with InitCons (32), MetConsσ (or MidConsσ) ceases to have effect if speakers overlook the morphological relationship between base and derivative. Thus *paisatge* [pəj.ˈza͡d.ʒə] 'landscape' is standard in Catalan, even though its relationship to *país* [pə.ˈis] 'country' is not hard to detect. *Traïdor* 'traitor' is usually pronounced [trəj.ˈðo] despite the relation to *trair* [trə.ˈi] 'betray'. The variant [mə.təw.ɾu.ˈlɔ.ʒik] for *meteorològic* perhaps not unreasonably ignores the connection with *meteor*. In faster informal speech, MetConsσ may be ineffective even in inflexional paradigms so that *enraonar* may be

pronounced [ən.rəw.ˈna] rather than [ən.rə.u.ˈna] and *amoïnar* [ə.muj.ˈna] (or [ə.mwi.ˈna] by FPH) rather than [ə.mu.i.ˈna]. It is likely that in all such examples the rhythmic constraint *LAPSE comes into play to drive glide formation in longer words. Recasens (1993: 110) observes that diphthongization is more likely in *enraonar* and *amoïnar* where the diphthong is internal than in *deïficar* 'deify.INF' or *laïcitzar* 'laicize.INF',[11] where it would be in the initial syllable. It is not surprising that in colloquial speech, purely phonological constraints such as ONSET and *LAPSE3σ are given more weight (i.e. are ranked higher) than morphological alignment or morphological correspondence constraints. The pronunciations mentioned in (34b), such as *enraonar* [ən.rə.u.ˈna], *arqueològic* [ər.ke.u.ˈlɔ.ʒik], reflect the ranking METCONSσ » *LAPSE3σ. The alternative ranking generates diphthongs: [pɾu.βəj.ˈðo], [ən.rəw.ˈna], [bəj.ku.ˈlar], [ə.muj.ˈna] ~ [ə.mwi.ˈna], [u.βəj.ˈɾi.əm], [ər.kew.ˈlɔ.ʒik], [mə.təw.ɾu.ˈlɔ.ʒik].

There remains a residue of cases, however, where hiatus is found (sometimes in variation with a falling diphthong) despite the absence of morphological justification 'protecting' the high vowel from glide formation. Examples are: *aorist* [ə.u.ˈɾist] 'aorist', *pleonasme* [plə.u.ˈnaz.mə] 'pleonasm', *Teodora* [te.u.ˈðo.ɾə] 'Theodora', *caïnita* [kə.i.ˈni.tə] 'kainite', *cuneïforme* [ku.nə.i.ˈfor.mə] 'cuneiform', *pioner* [pi.u.ˈne] 'pioneer'. There seems little alternative but to suppose that such words carry a lexical mark of syllabicity in their glide-resistant /i/ or /u/. That is to say, words such as these offer further evidence of irreducible, though very limited, contrast between high vowels and glides (see §3.2.3.2 above). Their resistance to glide formation is regarded here as faithfulness. As Serra (1996b) points out, it is more appropriate to mark the syllabicity as exceptional, inasmuch as a glide after another vowel would be the unmarked outcome.

3.2.4 *High vocoids preceding other vocoids*

I turn now to consider the alternative possibilities of rising diphthong or hiatus when a high vocoid precedes another vocoid. A major issue for how these sequences are syllabified in Catalan is variation, both of dialect and register. In comparison with the cases considered in §3.2.3, in the case of iV and uV sequences, morphological correspondence constraints have been found to be of less importance and prosodic constraints of greater importance (Cabré & Prieto 2004). I have suggested above (§3.2.2) that except when lexically marked for stress (*tia* [ˈti.ə] 'aunt', *pua* [ˈpu.ə] 'point'), unmarked syllabification imposes glide formation (*iaia* [ˈjajə] 'granny', *cauen* [ˈkawən] 'fall.3PL.PRS.IND'); but this

[11] It is DERIVCONSσ rather than METCONSσ that is involved in *laïcitzar*, and doubtless DERIVCONSσ (29b) should include in its scope the verb-forming affix /+ifik/ seen in *deïficar* and the abstract noun-forming suffix /+itat/ seen in *espontaneïtat*. Other initial syllable examples Recasens mentions that are not inclined to glide formation, such as *fluïdesa* 'fluidity', *coïssor* 'stinging pain', *lluïssor* 'brightness', are morphologically complex and not subject to *LAPSE3σ, having only two pretonic syllables.

is itself constrained by *ComplexOnset, so *Tià* [ti.ˈa] 'Seb' (hypocoristic of *Sebastià* 'Sebastian') is preferred to *[ˈtja]. It is shown here that *ComplexOnset must be unpacked into several onset constraints permitting different types of complex onset in different circumstances.

3.2.4.1 *Labio-velar obstruents?*

The first point to make concerns 'labio-velar' sequences [kw] and [gw] ∼ [ɣw]. There is a considerable number of words in which [kwV] or [gwV] ∼ [ɣwV] rather than [kuV] or [guV] ∼ [ɣuV] is found categorically in all varieties and styles (37). Invariant [kw] is consistently represented orthographically by *qu* before *a* or *o* and by *qü* before *e* or *i*. Variable [ku] ∼ [kw] is represented orthographically as *cu* or *co*.

(37) *quan* [ˈkwan] 'when'
Pasqua [ˈpas.kwə] 'Easter'
obliqua [u.ˈβli.kwə] 'oblique.F'
qüestió [kwəs.ˈtjo] 'question'
quota [ˈkwɔtə] 'quota'
guatlla [ˈgwaʎ.ʎə] 'quail'
aguantar [ə.ɣwən.ˈta] 'endure'
ungüent [uŋ.ˈgwen] 'ointment'
ambigua [əm.ˈbi.ɣwə] 'ambiguous.F'

In words with a morpheme boundary between the high and non-high vocoid, hiatus may be found after a velar obstruent, reflecting alignment or correspondence constraints of the type already mentioned in §3.2.3; variable glide formation is not excluded in (38).

(38) *cuota* [ku.ˈɔtə] 'tail.PEJ', *cuera* [ku.ˈerə] 'wagtail' (cf. *cua* [ˈku.ə] 'tail'—many derivatives of *cua* show the same pattern)
coacció [ku.ək.ˈsjo] 'coercion' (likewise other words containing the prefix *co-*)
coent [ku.ˈen] 'stinging, smarting' (stem /kɔ-/ seen in *coure* 'burn, sting')
evacuar [ə.βə.ku.ˈa] 'evacuate.INF' (cf. *evacua* [ə.βə.ˈku.ə] 'evacuate. 3SG.PRS.IND'; [ə.βə.ˈkwa] is also possible; see §3.2.4.7)
the feminine adjectives which correspond to masculine adjectives in *-cu* [-ku] (glides are possible here: see §3.2.4.3):

conspícua 'conspicuous'
innòcua 'innocuous'
perspícua 'perspicuous'
profícua 'profitable'
promíscua 'promiscuous'
vàcua 'vacant'
egoista [ə.ɣu.ˈis.tə] 'selfish', cf. *ego* 'ego'
goès [gu.ˈɛs] 'Goan', cf. *Goa* 'Goa'

In a small number of words with no morpheme boundary after the high vocoid, hiatus may be found, possibly in variation with a rising diphthong (39).

(39) *alcohol* 'alcohol'
coagular 'congeal'
coala 'koala'
coartar 'curtail'
coet 'rocket'
coercir 'force'
cohort 'cohort'

argüir [əɾ.ɣu.ˈi] 'imply' (/əRgu+i+R/—the verb *enaiguar* [ə.nəj.ˈɣwa] 'flood' shows that a following verb inflection is not of itself sufficient to block glide formation)
goècia 'black magic' (and *goètic* 'relating to black magic')

Invariant [kw], [gw] ~ [ɣw] as in (37), as opposed to the other types (38) and (39), has been attributed in some previous phonological analyses (Lleó 1970; Wheeler 1975/9; 1977) to the presence in the former of underlying unit phonemes /k^{w}/, /g^{w}/, while the others have underlying /ku/, /gu/. Broadly, this reflects the etymological situation and makes sense of the invariant presence of post-consonantal glides in this case in contrast with variable degrees of resistance to post-consonantal glide formation in all other contexts. But there are problems, even from a more traditional GP perspective. After unit obstruent labio-velar phonemes /k^{w}/, /g^{w}/, we might expect to find similar permitted consonants to what we find after /p/, /k/, /b/, /g/. That is, inasmuch as /pl/, /kl/, /bl/, /gl/, /pɾ/, /kɾ/, /bɾ/, /gɾ/ are permissible onsets, so also should */k^{w}l/, */g^{w}l/, */k^{w}ɾ/, */g^{w}ɾ/ be, but they are not found. We might also expect to find /k^{w}/, /g^{w}/ after the stress in proparoxytone words, where other single obstruents are permitted. Just as we have words like *òliba* [ˈɔ.li.βə] 'owl', *àguila* [ˈa.ɣi.lə] 'eagle', we might expect to meet words of the form *[ˈɔ.li.ɣwə] (/ˈɔlig^{w}+a/), *[ˈa.kwi.lə] (/ˈak^{w}il+a/), but we do not (with just one exception: *alíquota* [ə.ˈli.kwo.tə] 'aliquot'). From an OT perspective there is an additional difficulty of principle: positing 'abstract' underlying /k^{w}/, /g^{w}/ is a purely representational solution to a difference which is not in phonetic substance but in lexical distribution. The speaker/learner can observe invariant [kw], [gw], alongside variable [kw] ~ [ku], [gw] ~ [gu]. The evidence is sufficient to conclude that [w] is necessarily part of the onset in the former case. I interpret this here as a matter of faithfulness to underlying /w/, while in the case of variable [kw] ~ [ku], [gw] ~ [gu], the syllabic pronunciation will arise from correspondence constraints dominating *Cw, while the glide pronunciation will follow from the dominance of rhythmic or other markedness constraints.

The appeal previously to *ComplexOnset to penalize rising diphthongs involved an oversimplification. More precisely, what is involved is a constraint linking complexity of clusters with the markedness of glides in margins, namely *M/V$_{[+high]}$, given at (3). This complex onset constraint *ComplexM/V$_{[+high]}$, which ranks above *M/ V$_{[+high]}$ because it is inherently more specific, is spelt out in (40).

(40) *ComplexM/V$_{[+high]}$: Glides are not found in complex margins (*.Cj, *.Cw, *jC., *wC.).

Observe that (40) also penalizes complex codas with glides. Such codas are very rare in Catalan before internal onsets. There are a few examples like *claustre* 'cloister' or *augment* 'increase'. Morpheme-finally they are somewhat more common, as in *caut* 'cautious' or *Reus* (toponym); here a stressed diphthong follows in any case from the ranking *ˈV.V$_{[+high]}$ » *ComplexM/V$_{[+high]}$ (see

(26) and (28)). Complex codas of this type do indeed arise in concatenated morphemes or words for rhythmic reasons, as in *coincidir* [kujn.si.ˈði] 'to coincide', with a diphthong respecting *LAPSE3σ.

3.2.4.2 *Other post-consonantal high vocoids followed by non-high vocoids*

The remaining cases of post-consonantal high vocoids preceding other vocoids, such as in (41), are subject to some degree of variation.

(41)

diagonal 'diagonal'	*pianista* 'pianist'
dualisme 'dualism'	*piano* 'piano'
genuí 'genuine'	*quiet* 'still'
juliol 'July'	*suau* 'gentle'
llàntia 'lamp'	*suavitat* 'gentleness'
nació 'nation'	*variable* 'variable'
	vídua 'widow'

Some varieties, such as Balearic (Dols 2000 and p.c.), allow no post-consonantal rising diphthongs beyond those mentioned in §3.2.4.1. This is also the tradition of normative grammar—more specifically, of normative verse prosody (Cardona 1977). In spontaneous continental Catalan, rising diphthongs are found variably, with higher proportions among urban speakers, among younger speakers, and in faster or less formal speech (Cabré & Prieto 2004; Recasens 1993: 103–21). These circumstances are indicative of linguistic change in progress. From the OT perspective, this process of change also reflects the 'emergence of the unmarked' in respect of the syllabification of high vocoids. The general ONSET constraints are being promoted above the complex margin constraint *COMPLEXM/V$_{[+high]}$ (40). Within this variation, pronunciations with [j] rising diphthongs are in advance of those with [w] rising diphthongs. It is not immediately apparent why this should be so, but two related facts may be pointed out. First, word-initial [w] occurs only in recent borrowings and neologisms (such as *whisky* [ˈwiski], *UEFA* [ˈwɛfə] 'UEFA', *web* [ˈwɛp] 'world-wide web', *UAB* [ˈwap] (acronym of *Universitat Autònoma de Barcelona*)), while word-initial [j] is better integrated into the lexicon (as in *hiena* [ˈjɛnə] 'hyena', *iogurt* [juˈɣurt]). Secondly, the number of words with potential [j] diphthongs is considerably greater than that of words with potential [w] diphthongs, so once the change has got under way, speakers/learners are presented with more frequent evidence for [Cj] than for [Cw].

Here I investigate the various phonological factors which favour or hinder the formation of rising diphthongs. Recent studies have come to different conclusions about the nature of these factors. For example, Recasens (1993: 103–21), Bonet & Lloret (1998: 64), and Jiménez (1999: 69) all claim that the presence of an already complex onset is or may be a barrier to glide formation. They report hiatus in words such as *síndria* [ˈsin.dɾi.ə] 'watermelon', *indústria* [in.ˈdus.tɾi.ə] 'industry', *embrió* [əm.bɾi.ˈo] 'embryo', and *ampliar* [əm.pli.ˈa] 'expand'. Such a constraint is entirely plausible, since triply filled onsets are not generally found in Catalan—elsewhere there is evidence for a *CCCONSET constraint. Cabré & Prieto (2001; 2004), however, on the basis of a questionnaire study of sixty informants, find no evidence that words with this syllable make-up are treated

differently from words involving single consonant onsets. Recasens (1993) and Jiménez (1999: 75) also claim that output–output correspondence effects propagate diphthongs or hiatuses from bases to derivatives; for example *embrionari* [əm.bɾi.u.ˈna.ɾi] 'embryonic' after *embrió* [əm.bɾi.ˈo] 'embryo' (Recasens), *aliança* [a.li.ˈan.sa] 'alliance' after *aliar* [a.li.ˈa] 'to ally' (Jiménez's transcriptions represent Valencian pronunciation). Cabré & Prieto's study again finds no evidence of correspondence effects when the alternatives are hiatus and rising diphthongs, though they are ready to acknowledge such effects in vocoid sequences of falling sonority. (This difference, we shall see, reflects the fact that output–output correspondence is nearly always outranked by other markedness constraints, when the alternatives under consideration are hiatuses and rising diphthongs.) I largely follow Cabré & Prieto's observations here, though the explanation I offer is different in certain respects.

Those varieties, such as Balearic and the variety of normative prosodics, that present no post-consonantal rising diphthongs display the constraint ranking *ComplexM/V$_{[+high]}$ » Onset (taking Onset as representative of constraints barring hiatus). The opposite ranking, Onset » *ComplexM/V$_{[+high]}$, would reflect the grammar of a variety with rising diphthongs in all possible contexts (provided no other relevant constraints outranked Onset). The dialects considered here, in which rising diphthongs are found in some contexts but not others, illustrate the demotion of some elements of *ComplexM/V$_{[+high]}$ below Onset, or the ranking of other glide-promoting constraints above *ComplexM/V$_{[+high]}$. The presence of [j] diphthongs alongside [u] hiatuses in similar contexts reflects the splitting of *ComplexM/V$_{[+high]}$, with *Cj being demoted further than *Cw.

3.2.4.3 *Vocoid sequences of rising sonority in relation to word stress*

I consider in turn vocoid sequences of rising sonority in post-tonic, word-initial, tonic, and pre-tonic position. In post-tonic position it is agreed by all observers that diphthongs are favoured in most continental varieties; diphthongs with [j] are even more widely favoured than those with [w] (42).

(42) *síndria* [ˈsin.dɾjə] 'watermelon' *pèrdua* [ˈpɛɾ.ðwə] 'loss'
llàntia [ˈʎan.tjə] 'lamp' *tènue* [ˈtɛ.nwə] 'tenuous'
sèrie [ˈsɛ.ɾjə] 'series' *vídua* [ˈbi.ðwə] 'widow'
família [fə.ˈmi.ljə] 'family'

One might say that pattern (42) merely reflects the simple case of ranking Onset above *ComplexM/V$_{[+high]}$; but as I shall shortly come to revise that general ranking, one might ask what additional reason there might be for favouring diphthongs particularly in this environment. In a preliminary version of Cabré & Prieto (2004), Cabré & Prieto (2001: 15)[12] plausibly suggest an alignment

[12] Cabré and Prieto do not carry this idea forward to the revised version of the paper (2004) because they do not believe output–output correspondence constraints interact with constraints involving rising diphthongs. From this perspective, without MetConσ in (44), AlignR[ˈF] can be dispensed with. They are left without an account of the hiatus pattern in (38). This hiatus may be variable, but their ranking *C$_{velar}$uV » Onset rules it out altogether.

constraint ALIGNR[ˈσσ] favouring the alignment of a stressed bisyllabic foot at the right edge of a word. Actually it is not essential that the foot referred to in the constraint be bisyllabic. In what follows I assume moraic FTBIN; a mora that cannot be parsed in a bimoraic foot is unparsed. But in any case, the constraint needs only to favour better-aligned feet of any type over worse-aligned feet. That is to say, the constraint would penalize all unstressed syllables at the right edge. It is formulated here as (43).

(43) ALIGNR[ˈF]: ALIGN(Stressed foot, R, PrWd, R).

The constraint ALIGNR[ˈF] appears to dominate metrical correspondence constraints in sets such as *vària* [ˈba.ɾjə] 'varied.F' – *varia* [bə.ˈɾi.ə] 'vary. 3SG.PRS.IND', *contínua* [kunˈti.nwə] 'continuous.F' – *continua* [kunti.ˈnu.ə] 'continue.3SG.PRS.IND'. In these examples a post-tonic onset glide [j] or [w] corresponds to a stressed vowel in a derivationally related word. And this overriding of output–output correspondence is one reason not simply to demote *COMPLEXM/$V_{[+high]}$ below ONSET. Another advantage of appealing to ALIGNR[ˈF] here is that variability is readily associated with rhythmic constraints, as can be seen with the variable effects of the *LAPSE constraints. Tableau (44) displays candidates for (a) *síndria* [ˈsin.dɾjə] 'watermelon', (b) *vídua* [ˈbi.ðwə] 'widow', and (c) *vària* [ˈba.ɾjə] 'varied.F'. The footing is iambic in accordance with the discussion in Chapter 9, but the analysis would work just as well if feet were trochaic. Alignment violations are counted in syllables. Intermediate varieties that prefer *contínua* [kunˈti.nu.ə] with hiatus, but have *vària* [ˈba.ɾjə] with a diphthong, have *Cw » ALIGNR[ˈF].

(44)

	síndria	ALIGNR[ˈF]	METCONSσ	ONSET	*COMPLEXM/$V_{[+high]}$
(a)	(ˈσ) (σσ)] ˈsin.dɾi.ə	σ!σ		*	
	(ˈσ) σ] ☞ ˈsin.dɾjə	σ			*
(b)	*vídua*				
	(ˈσ)(σσ)] ˈbi.ðu.ə	σ!σ		*	
	(ˈσ) σ] ☞ ˈbi.ðwə	σ			*
(c)	*vària* cf.*varia* [bəˈɾi.ə]				
	(ˈσ)(σσ)] ˈba.ɾi.ə	σ!σ		*	
	(ˈσ) σ] ☞ ˈba.ɾjə	σ	*		*

In the case of post-tonic [iu] sequences, such as in *pòdium* 'podium', *Màrius* (given name), the constraint ranking in (44) will favour diphthong formation over hiatus. Recasens remarks (1993: 107, 111) that it is hard to be certain whether the output is actually [ju] or [iw] in such cases. The constraint *ComplexM/V$_{[+high]}$ fails to choose between them, since each involves one glide in a complex margin. Formally, any choice falls to lower ranked constraints.

3.2.4.4 *High vocoid followed by stressed non-high vocoid in initial syllable*

I consider next word-initial post-consonantal vocoid sequences of rising sonority where the more sonorous vocoid carries lexical stress (45). In this context a strong preference for hiatus is found even among those speakers and styles at the forefront of change (Cabré & Prieto 2001; 2004).

(45)

viola [bi.ˈɔ.lə] 'violet'	*suor* [su.ˈo] 'sweat'
miop [mi.ˈɔp] 'short-sighted'	*suau* [su.ˈaw] 'gentle'
piano [pi.ˈa.nu] 'piano'	*lluert* [ʎu.ˈɛrt] 'lizard'
niar [ni.ˈa] 'deny.INF'	*pouet* [pu.ˈɛt] 'well.DIM'/ˈpou+ˈɛt/
fiança [fi.ˈan.sə] 'security'	*jueva* [ʒu.ˈe.βə] 'Jewish.F'
client [kli.ˈen] 'customer'	*lloança* [ʎu.ˈan.sə] 'praise'

The same pattern is evident in hypocoristic forms derived by truncation of given names that themselves contain a rising diphthong (Cabré & Prieto 2001: 5; Cabré 1998: 15 ff.), as in (46).

(46)

Base	Truncated hypocoristic
Sebastià [sə.βəs.ˈtja]	*Tià* [ti.ˈa]
Concepció [kun.səp.ˈsjo]	*Ció* [si.ˈo]
Daniel [də.ˈnjɛl]	*Niel* [ni.ˈɛl]

Why a language might preferentially avoid rising diphthongs in initial syllables is a question that needs some further investigation. It is not at first sight clear whether markedness or faithfulness is at issue. From the faithfulness perspective, one possibility relates to the prominence of word-initial position; diphthong avoidance would be a matter of 'positional faithfulness' (Beckman 1998). This is the approach followed by Cabré & Prieto (2004 a different, less successful appeal to Root Binarity is made in Cabré & Prieto 2001). The constraint they invoke is MaxInitμ: 'Every word-initial [syllable] input mora has an output correspondent.' For this to work requires moras to be underlying in initial-syllable high vocoids. Up to now in this chapter I have suggested that the syllabification, and implicitly the moraicity, of high vocoids is determined by general principles of markedness in syllabification, namely SonSeq and Onset constraints. The moraicity of high vocoids does, however, follow from the peak/margin constraint *M/V$_{[+high]}$ (3), which has the consequence that high vocoids are moraic—more specifically, bear head moras—unless some higher constraint outranks *M/V$_{[+high]}$. We have already observed, however, that Onset does so. Given 'Richness of the Base', we should seek a grammar in which the unmarked

or default pattern of initial syllable hiatus as in (45) is the one that follows from markedness considerations. This is also the most natural inference to be drawn from the derivational truncation pattern illustrated in (46), in which it is clear that some constraint outranks the correspondence constraint UNIFORMσ by which the derived form would be expected to match the base as far as possible in syllabification. It is natural to see deviation from exact syllabic correspondence in the derived truncated names as showing an 'emergence of the unmarked' effect, and this is implicitly the claim of Cabré (1998), where hiatus in truncated hypocoristics follows from a version of *COMPLEXONSET.

From the markedness perspective, it might well be that the avoidance of rising diphthongs in initial syllables is a question of avoiding initial complex onsets. I mentioned *COMPLEXONSET at (19) and gave a more specific version, *COMPLEXM/V$_{[+high]}$, at (40). Then the pattern in (45), (46), and (52) would reflect a positional version of *COMPLEXM/V$_{[+high]}$, limited to initial syllables. The insight that led to Cabré & Prieto's MAXINITμ as an example of positional faithfulness can thus be recast as positional markedness. While the general markedness constraint against glides, *M/V$_{[+high]}$, stands below ONSET, one might suggest a version restricted to word-initial syllables that in turn ranks above ONSET. To reflect the default character of the initial hiatus pattern in (45), then, Cabré & Prieto's MAXINITμ should be replaced by a markedness version as in (47).

(47) *INITIALM/V: The first vocoid in a lexical form is not a margin.

A more controversial proposal would be to relate hiatus after initial-syllable high vowels to the claim made in Chapter 9, for other reasons, that stress in Catalan is preferentially iambic. Initial pretonic hiatus supports a well-formed bisyllabic iambic foot (σˈσ) in *miop* [miˈɔp] as opposed to a degenerate iamb (ˈσ) in *[ˈmjɔp]. (The head would be bimoraic in either case.) This approach would not extend automatically to the initial pretonic cases of hiatus in (52), however.

Native speakers agree that there are a few words with initial post-consonantal rising diphthongs even for those speakers whose speech generally corresponds to the situation illustrated in (45) and (46). There is less agreement on which these words are. In (48) I list the particular cases mentioned by Recasens (1993: 114, 118) and Cabré & Prieto (2001: 9; 2004). Only five of the eleven are common to both sets of authors, and there is explicit disagreement about *suec* 'Swedish'.

(48)

Words with initial stressed rising diphthong	Recasens	Cabré & Prieto
fuel [ˈfwɛl] 'fuel oil'		✓
suèter [ˈswɛ.tər] 'sweater'		✓
suec [ˈswɛk] 'Swedish'	×	✓
duel [ˈdwɛl] 'duel'	✓	✓
ciència [ˈsjɛn.sjə] 'science'	✓	✓
Viena [ˈbjɛ.nə] 'Vienna'	✓	✓
quiet [ˈkjet] 'calm'	✓	✓

triomf [ˈtɾjomɟf] 'triumph'	✓	✓
Siena [ˈsjɛ.nə] 'Siena'	✓	
dièresi [ˈdjɛ.ɾə.zi] 'diaeresis'	✓	
diòcesi [ˈdjɔ.sə.zi] 'diocese'	✓	

This state of affairs resembles the initial stage of lexical diffusion in a change in progress. For the speakers using the diphthong in these (or other) words, there is an incipient lexical contrast between vowel and glide in this position (so these words will contain lexical /w/ or /j/ (non-moraic).[13] There is no doubt an element of linguistic interference involved, as both Recasens and Cabré & Prieto remark. Most of the words in (48) have close phonological cognates in Spanish, containing cases of the stressed rising diphthongs [jɛ] and [wɛ] which are very common in that language. Catalan words such as *coet* [ku.ˈɛt] 'rocket' and *poeta* [pu.ˈɛtə], despite having a high round + low mid front sequence, avoid the diphthong realization [wɛ], perhaps because the corresponding words in Spanish have no high vowel and no diphthong (*cohete* [ko.ˈete] and *poeta* [po.ˈeta]).

Cabré & Prieto's account of the pattern in (45), based on MaxInitμ, involves the invocation of two further constraints. One of these is to enforce 'labio-velar' clusters [kw], [gw] ~ [ɣw], as seen in (37) over initial syllable hiatus, and indeed over hiatus anywhere involving [ku.V] or [gu.V] ~ [ɣu.V]. The constraint is simply *C_{velar}uV, for which Cabré & Prieto do not offer functional justification. Further research is needed on the greater acceptability of 'labio-velar' clusters. But a problem with invocation of *C_{velar}uV, outranking MaxInitμ, is that it fails to reflect any distinction between the categorical 'labio-velar clusters' of (37) and the variable ones of (38), where the variability is parallel to all the other cases of [Cu.V] ~ [CwV] considered in §3.2.4. In Cabré & Prieto's grammar, the pronunciation [kunsˈpiku.ə] for *conspícua* 'conspicuous.F' or [ku.ˈɔtə] for *cuota* 'tail.PEJ' is impossible. So, as suggested above, I take categorical 'labio-velar' clusters to reflect the dominance of faithfulness (Ident[w]) over markedness constraints that might promote hiatus. The other constraint Cabré & Prieto need (2004), given MaxInitμ/*InitialM/V, is the constraint that will favour, for *iuca* 'yucca' (19), the pronunciation [ˈjukə] over *[i.ˈukə], and similarly for all word-initial high vocoid+vocoid sequences. The constraint they propose is Onset+Onset, which penalizes two consecutive Onset violations. That is, the constraint penalizes *.V.V—in this case *[i.ˈu.kə]. This analysis cannot be right, however. With the ranking Onset+Onset » MaxInitμ » Onset a candidate with a falling diphthong, namely *[ˈiw.kə], would win. Its first vocoid is moraic (as is its second). Only Onset is violated. The correct winner, [ˈju.kə], violates higher ranked MaxInitμ. The required constraint of the Onset family might be InitialOnset (49), which naturally is outranked by Max and Dep, which could achieve the initial onset by deletion or insertion.

[13] Or one might invoke co-phonologies (Inkelas et al. 1997); the 'exceptions' grammar would have Onset » *InitialM/V.

(49) InitialOnset (InitOns): The initial syllable of every prosodic word has an onset.

However, there is a more transparent way of interpreting the situation which called for InitialOnset » MaxInitμ/*InitialM/V, which is to identify simply *ComplexM/V$_{[+high]}$, relativized by 'positional markedness' to initial position, as stated in (50).

(50) *InitialComplexM/V$_{[+high]}$ (*InCompM/V): A glide is not found in a complex margin in the initial syllable of a prosodic word.[14]

The account of post-consonantal word-initial vocoid sequences of rising sonority that I adopt here reflects the observation that *InitialComplexM/V$_{[+high]}$ generally outranks Onset, but is outranked by a faithfulness constraint for underlying /j, w/ in the examples of (37) and (48). Tableau (51) illustrates the alternative candidates for (a) *miop* [mi.ˈɔp] 'short-sighted' (b) *quiet* [ˈkjet] 'calm', (c) *iuca* 'yucca', (d) *quota* 'quota', and (e) *cuota* 'tail.pej' (cf. *cua* [ˈku.ə] 'tail').

(51)

		Id/j,w/	*InCompM/V	Onset	*ComplexM/V	*M/V$_{[+high]}$
(a)	*miop* /miɔp/					
	ˈmjɔp		*!		*	*
	☞ mi.ˈɔp			*		
(b)	*quiet* /kjet/					
	☞ ˈkjet		*		*	*
	ki.ˈet	*!		*		
(c)	*iuca*					
	i.ˈu.kə			**!		
	☞ ˈju.kə					*
	ˈiw.kə			*	*!	*
(d)	*quota* /ˈkwɔta/					
	☞ ˈkwɔ.tə		*		*	*
	ku.ˈɔ.tə	*!		*		
(e)	*cuota* cf. *cua*					
	ˈkwɔ.tə		*!		*	*
	☞ ku.ˈɔ.tə			*		

Id/j,w/ » *InCompM/V » Onset, *ComplexM/V$_{[+high]}$ » *M/V$_{[+high]}$

[14] The *InitialComplexM/ V$_{[+high]}$ constraint is similar but not identical to the one proposed by Jiménez (1999: 72) in this context. Jiménez, without generally arguing for iambic stress, invokes a

3.2.4.5 *Word-initial pretonic sequences of rising sonority*

The analysis developed in §3.2.4.4 can be carried forward with no modification to cases of post-consonantal word-initial vocoid sequences of rising sonority that are unstressed but immediately pretonic. *INITIALCOMPLEX M/V$_{[+high]}$ as formulated in (47) makes no mention of stress (unlike the version proposed by Jiménez (1999: 72)). Here again hiatus is the norm (Recasens 1993: 118–19; Cabré & Prieto 2001: 10; 2004). Examples are given in (52). Morphological structure is irrelevant to the question of hiatus versus rising diphthong.

(52) *dialecte* [di.ə.ˈlek.tə] 'dialect'
dialoga [di.əˈ.lɔ. ɣə] 'engage-in-dialogue.3SG.PRS.IND. '
pianista [pi.ə.ˈnis.tə] 'pianist'
violí [bi.u.ˈli] 'violin'
biològic [bi.u.ˈlɔ.ʒik] 'biological'
cianur [si.ə.ˈnur] 'cyanide'
viatger [bi.ə͡dˈ.d͡ʒe] 'traveller'
riallada [ri.ə.ˈʎa.ðə] 'guffaw'
dualisme [du.ə.ˈliz.mə] 'dualism'
Joaquim [ʒu.ə.ˈkim] (given name)
suaria [su.ə.ˈɾi.ə] 'sweat.3SG.COND'
cuereta [ku.ə.ˈɾɛ.tə] 'wagtail'
jueria [ʒu.ə.ˈɾi.ə] 'Jewish quarter'

Again, there are a few words mentioned as exceptions. Cabré & Prieto cite *pietat* [pjəˈtat] 'pity' (though for Recasens this is [pi.əˈtat]); Recasens mentions that items from (48) retain rising diphthongs when suffixed, giving *científic* [sjən.ˈti.fik] 'scientific', *triomfal* [tɾjuɱ.ˈfal] 'triumphal' (as is expected if /j/ is lexical in such words).

3.2.4.6 *Initial pre-pretonic sequences of rising sonority*

The pattern of pronunciation of longer words with potential rising diphthongs in initial syllables differs from what was seen in §§3.2.4.4–3.2.4.5. When at least one syllable intervenes between the potential initial diphthong and the stressed syllable, a diphthong is preferred (53). In (53b) some paradigmatically related sets illustrate the systematicity of the contrast between pre-pretonic (or pre-pre-pretonic) on the one hand and pretonic and tonic on the other.

(53) (a) *diagonal* [djə.ɣu.ˈnal] 'diagonal'
diapasó [djə.pə.ˈzo] 'scale'
diapositiva [djə.pu.zi.ˈti. βə] 'slide'
violoncel [bju.lun.ˈsɛl] 'cello'
suavitat [swə.βi.ˈtat] 'gentleness'

(b)

3 before tonic	2 before tonic	1 before tonic	tonic
dialogaré	*dialogar*	*dialoga*	*diàleg*
[d**jə**.lu.ɣə.ˈɾe] 'engage-in-dialogue.1SG.FUT'	[d**jə**.lu.ˈɣa] 'engage-in-dialogue.INF'	[d**i.ə**.ˈlɔ.ɣə] 'engage-in-dialogue. 3SG.PRS.IND'	[d**i.**ˈ**a**.lək] 'dialogue'

NON-INITIAL stress constraint, which he restricts to syllables with rising diphthongs. But his constraint leaves out of account hiatus after an initial unstressed syllable, as in (52) below.

diabolical	*diabolisme*	*diabòlic*	*diable*
[d**jə**.βu.li.ˈkal]	[d**jə**.βu.	[d**i.ə**.ˈβɔ.lik]	[d**i**.ˈ**a**b.blə]
'diabolical'	ˈliz.mə]	'diabolical'	'devil'
	'satanism'		
	biologia	*biològic*	*biòleg*
	[b**ju**.lu.ˈʒi.ə]	[b**i.u**.ˈlɔ.ʒik]	[b**i**.ˈ**ɔ**.lək]
	'biology'	'biological'	'biologist'
violoncel·lista	*violinista*	*violí*	*viola*
[b**ju**.lun.sə.ˈlis.tə]	[b**ju**.li.ˈnis.tə]	[b**i.u.**ˈli]	[b**i**.ˈ**ɔ**.lə]
'cellist'	'violinist'	'violin'	'viola'

In §3.2.3.2 I drew attention to a rhythmic constraint, *LAPSE3σ, disfavouring a succession of more than two unstressed syllables. I argued then that *LAPSE3σ is responsible for preferring *preocupant* 'worrying' [prəw.ku.ˈpan] to [prə.u.ku.ˈpan], while for *preocupa* 'concern.3SG.PRS.IND', [prə.u.ˈku.pə] with hiatus is preferred to [prəw.ˈku.pə]. It should now be evident that the situation I have been considering in §§3.2.4.4–3.2.4.6 is parallel in essential respects. Diphthongs are preferred over hiatus when the alternative is a sequence of more than two unstressed syllables. When a sequence of more than two unstressed syllables is not in the picture, of course *LAPSE has no role. With *LAPSE4σ inherently outranking *LAPSE3σ, the constraint ranking *LAPSE4σ » *LAPSE3σ, IDENT/j,w/ » METCONSσ » *INITIALCOMPLEX M/V$_{[+high]}$ » ONSET will account for the distribution of hiatus and rising diphthongs seen in §§3.2.4.4–3.2.4.6. The tableau (54) illustrates this with (a–c) *dialogaré*, *dialogar*[15] and *dialoga* taken from (53b), and with two alternative pronunciations of *pietat* 'pity': (d) [pi.ə.ˈtat], which follows the general pattern, and (e) [pjə.ˈtat] for speakers for whom this word is lexicalized with /j/.

3.2.4.7 *Non-initial sequences of rising sonority, with second element stressed*

Next I turn to post-consonantal non-initial vocoid sequences of rising sonority where the more sonorous element is stressed. Diphthongs are observed to be virtually universal in nouns containing the suffix *-ció* or *-sió*, irrespective of how many syllables precede: *estació* [əs.tə.ˈsjo]'station', *compressió* [kum.prə.ˈsjo] 'compression', *il·lusió* [i.lu.ˈzjo] 'dream', *nació* [nə.ˈsjo] 'nation', *fusió* [fu.ˈzjo] 'fusion'. Apart from words of the type just mentioned, words with the sequence high vocoid–stressed non-high vocoid preceded by one syllable provide the context in which there seems to be most variation between hiatus and a rising diphthong, even among speakers of the same variety. According to Recasens

[15] A reader observes that, with secondary stress, a candidate [di.ˌə.lu.ˈɣa] would beat [djə.lu.ˈɣa], since violations of *LAPSE3σ, METCONSσ, and *INITCOMPM/V would all be avoided. A provisional account of secondary stress is offered in §9.6. I observe that, as a 'repair' to a *LAPSE violation, glide formation is in fact preferred to secondary stressing, but I do not currently have an account of this observation. It may perhaps result from something like a *STRUCTURE constraint, such as is mentioned at the end of §4.4.

(54)

		*LAPSE4σ	*LAPSE3σ	ID/j,w/	METCONSσ	*INCOMPM/V	ONSET
(a)	*dialogaré*						
	σ σ σ 'σ ☞ dj ə.lu. ɣə.'ɾe		*		*	*	
	σ σ σ σ 'σ di. ə.lu. ɣə.'ɾe	*!	*				*
(b)	*dialogar*						
	σ σ 'σ ☞ dj ə.lu.'ɣa				*	*	
	σ σ σ 'σ di. ə.lu.'ɣa		*!				*
(c)	*dialoga*						
	σ 'σ σ dj ə.'lɔ.ɣə					*!	
	σ σ 'σ σ ☞ di. ə.'lɔ.ɣə						*
(d)	*pietat*						
	σ 'σ pj. ə.'tat					*!	
	σ σ 'σ ☞ pj. ə.'tat						*
(e)	*pietat* /pj ə.'tad/						
	σ 'σ ☞ pj ə.'tat					*	
	σ σ 'σ pi. ə.'tat			*!			*

(1993: 116–20) and Cabré & Prieto (2001; 2004), rising diphthongs predominate. Among the majority group of Cabré & Prieto's informants, whose dialect they call the 'General variety', rising diphthongs were preferred in words of the (55) type in more than two-thirds of cases. A minority among their informants were speakers of a 'Conservative variety', who preferred hiatus in words of the type in (55) (63 per cent for the paroxytone *moniato* type, 58 per cent for the oxytone *juliol* type). In (55) is a sample of words (from Cabré & Prieto 2001) illustrating stressed rising non-initial diphthongs, with one syllable preceding.

(55) *clariana* [klə.ˈɾja.nə] 'clearing'
moniato [mu.ˈnja.tu] 'sweet potato'
idioma [i.ˈðjo.mə] 'language'
juliol [ʒu.ˈljɔl] 'July'
aviat [ə.ˈβjat] 'soon'
radiòleg [rə.ˈðjɔ.lək] 'radiologist'
confiança [kuŋ.ˈfjan.sə] 'trust'
genuí [ʒə.ˈnwi] 'genuine'
manual [mə.ˈnwal] 'manual'

The two groups of authors present different lists of individual words that are most likely to show hiatus rather than a rising diphthong in this context even among speakers of the 'General' variety, though *atiar* [əti.ˈa] 'poke a fire' and *esquiar* [əski.ˈa] 'ski$_{v}$' are on both lists—in fact, for Cabré & Prieto all the hiatus examples are, like these two, verbs of conjugation I with bisyllabic stems.[16] Recasens has *enciam* [ənsi.ˈam] 'lettuce', *adient* [əði.ˈen] 'appropriate', *guardiola* [gwərði.ˈɔlə] 'money box', *Damià* [dəmi.ˈa] (given name), for all of which Cabré & Prieto report rising diphthongs. These authors, in fact, say (2001: 3): 'for a given speaker, a lexical item shows either diphthong or hiatus, but not both, and in some isolated cases both diphthong and hiatus are possible.' That is, the contrast is basically lexical, but change is in progress, with the trend being to transfer words to the diphthong class, which seems now to be unmarked in this context.

This state of affairs is what is reflected in the constraint ranking ONSET, *COMPLEXM/V$_{[+high]}$. As has been shown in (51), *INITIALCOMPLEXM/V$_{[+high]}$ outranks ONSET so the constraint *COMPLEXM/V$_{[+high]}$ (expressed in general terms) reflects the 'elsewhere' case. It will only serve to evaluate candidates for which the more specific constraint *INITIALCOMPLEXM/V$_{[+high]}$ is not relevant. Consequently, words with hiatus in this pretonic non-initial context will be the ones with lexical marking, for which IDENT/i,u/ will outrank ONSET. Insofar as the words belonging to the hiatus class are conjugation I verbs, lexical marking need not be involved, as the correspondence constraint (35) might be invoked: METCONSσ (The head of a syllable in a base corresponds to the head of a syllable in a derivative). Observe, for example, that *esquiar* [əs.ki.ˈa] 'ski.INF' corresponds to the base form *esquia* [əs.ˈki.ə] 'ski.3SG.PRS.IND'. In fact, though, given that we have previously seen correspondence constraints outranking ONSET constraints (§§3.2.3.2–3.2.3.3) so as to prevent glide formation in the case of falling diphthongs, the issue to consider involves words like *variable* [bə.ˈɾjab.blə] 'variable' which corresponds to *varia* [bə.ˈɾi.ə] 'vary.3SG.PRS.IND', or *confiança* [kuŋ.ˈfjan.sə] 'trust' which corresponds to *confia* [kuŋ.ˈfi.ə] 'trust. 3SG.PRS.IND'. This suggests rather that a stronger diphthong constraint is outranking correspondence for the *confiança* type, while it is indeed IDENT/i,u/ that preserves hiatus in specific words of the *esquiar* type for specific speakers. Why might [i.ˈa] be less harmonic than [ˈja], to a greater degree than [ə.i] is less

[16] The remaining items on their list are: *arriar* 'tow', *destriar* 'separate', *enviar* 'send', *espiar* 'spy', *graduar* 'grade'.

harmonic than [əj]? A constraint against hiatus between a stressed vowel and a high vocoid was introduced at (28): *ˈV.V$_{[+hi]}$: A stressed vowel is not immediately followed by a high vowel nucleus. The sequence that is avoided in *confiança* is the mirror image of that one. And the reason why the sequence V$_{[+hi]}$.ˈV might be avoided is likewise perceptual, I believe. An unstressed vowel is typically shorter than a stressed one. A short high vocoid followed by a longer, stressed vowel is likely to be *interpreted* as a rising diphthong, especially when the pattern of rising diphthongs is already established in the language. A high vocoid followed by an unstressed vowel is less likely to be (mis)perceived as a rising diphthong. The matching 'no stressed hiatus constraint' is, then, as in (56).

(56) *V$_{[+hi]}$.ˈV: A stressed vowel is not immediately preceded by a high vowel nucleus.

Tableau (57) illustrates the evaluation of candidates against the constraints mentioned in this section: (a) is *aviat* 'soon', (b) is *variable* 'variable', and (c) is *esquiar* 'ski.INF'

(57)

		IDENT/i,u/	V$_{[+hi]}$.ˈV	METCONSσ	ONSET	*COMPM/V
(a)	*aviat*					
	əβi.ˈat		*!		*	
	☞ əˈβjat					*
(b)	*variable* cf. bə.ˈri.ə					
	bə.ri.ˈab.blə		*!		*	
	☞ bə.ˈrjab.blə			*		*
(c)	*esquiar* cf. əs.ˈki.ə					
	☞ əs.ki.ˈa		*		*	
	əs.ˈkja	*!		*		*

IDENT/i,u/ » V$_{[+hi]}$.ˈV » METCONSσ » ONSET, *COMPLEXM/V$_{[+high]}$

Recasens (1993: 116) observes an interesting phenomenon among older speakers in the Camp de Tarragona for whom *COMPLEXM/V$_{[+high]}$ seems generally to be ranked above ONSET (and also above ALIGNR[ˈF]). He records *partió* [pərti.ˈo] 'boundary', *nerviós* [nərvi.ˈos] 'upset', but *injecció* [in^{j}d͡ʒək.ˈsjo] 'injection', *estalviar* [əstəl.ˈvja] 'save'. With the ranking *COMPLEXM/V$_{[+high]}$ » ONSET, *LAPSE again becomes relevant as it was in (54). In this variety the evaluation of candidates for *nerviós* and *estalviar* is as in (58).

(58) Camp de Tarragona, older speakers

nerviós	*LAPSE3σ	*COMPM/V	ONSET
nər.ˈvjos		*!	
☞ nɚvi.ˈos			*
estalviar			
☞ əstəl.ˈvja		*	
əstəlvi.ˈa	*!		*

The pattern observed in (58) among older speakers in the Camp de Tarragona differs only slightly from the pattern of Cabré & Prieto's 'conservative' variety. In this variety, words of the type in (55) have hiatus. Thus for speakers of this variety the constraint against rising diphthongs (*COMPLEXM/$V_{[+high]}$) is overridden only by (a) faithfulness, in the case of lexical 'labiovelar' clusters (37), (b) rhythmic constraints, namely *LAPSE3σ, and ALIGNR[ˈF], and whatever favours [-ˈsjo], [-ˈzjo] for the suffixes *-ció* and *-sió*.[17] Of these, *LAPSE3σ induces diphthongs to avoid three pretonic stressless syllables, and ALIGNR[ˈF] induces diphthongs to avoid two post-tonic stressless syllables. In fact, even in the more general variety the dominance of *LAPSE3σ over *$V_{[+hi]}$.ˈV and ONSET means that it is in fact the *LAPSE3σ violation, rather than the *$V_{[+hi]}$.ˈV violation, that causes [əstəlˈβja] to win over [əstəlβiˈa] for *estalviar*, or, indeed, [kumprəˈsjo] over [kumprəsiˈo] for *compressió*.

The same constraint set accounts generally for stressed rising diphthongs in words with more than one pretonic syllable, i.e. σσ.ˈCjV or σσ.ˈCwV, as exemplified in (59).

(59) *farmaciola* [fər.mə.ˈsjɔ.lə] 'first-aid kit'
Monturiol [mun.tu.ˈɾjɔl] (surname)
eclesiàstic [ə.klə.ˈzjas.tik] 'church$_A$'
negociar [nə.ɣu.ˈsja] 'negotiate'
entusiasme [ən.tu.ˈzjaz.mə] 'enthusiasm'
mobiliari [mu.βi.ˈlja.ɾi] 'household goods'
respectuós [rəs.pək.ˈtwos] 'respectful'
estatueta [əs.tə.ˈtwɛ.tə] 'statue.DIM'

3.2.4.8 *Pretonic non-initial sequences of rising sonority*

The final context to consider, for post-consonantal vocoid sequences that are potential rising diphthongs, is that which is pretonic but not initial. Diphthongs are the norm here, except in formal styles. In (60a) are some examples with immediately pretonic diphthongs; in (60b) some with pre-pretonic diphthongs.

[17] This could be DERIVATIONALCONSISTENCYσ. In the majority of lexemes these suffixes are preceded by stems of two or more syllables, and *LAPSE3σ favours the diphthong allomorph.

(60) (a) *biblioteca* [bi.βlju.ˈtɛ.kə] 'library'
cristianisme [kɾis.tjə.ˈniz.mə] 'Christianity'
Maquiavel [mə.kjə.ˈβɛl] 'Machiavelli'
persuadir [pər.swə.ˈði] 'persuade'
ritualista [ri.twə.ˈlis.tə] 'ritualist'
continuarem [kun.ti.nwə.ˈɾɛm] 'continue.1PL.FUT'
continuïtat [kun.ti.nwi.ˈtat] 'continuity'
virtuosisme [bir.two.ˈziz.mə] 'virtuosity'
(b) *labialitzar* [lə.βjə.li͡d.ˈd͡za] 'labialize'
patrioterisme [pə.tɾju.tə.ˈɾiz.mə] 'chauvinism'
actualitat [ək.twə.li.ˈtat] 'present'
septuagenari [səp.twə.ʒə.ˈna.ɾi] 'septuagenarian'

The fact that contrast is unusual in words of this pattern, as opposed to the context that was the main focus of §3.2.4.7 (stressed diphthong ~ hiatus with one pretonic syllable) in which variation is strongest, suggests, in line with the observation at the end of §3.2.4.7, that the critical constraint here is * LAPSE3σ, which dominates *COMPLEXM/V$_{[+high]}$ for all speakers for whom any constraint dominates *COMPLEXM/V$_{[\ +high]}$ (including those who have *COMPLEXM/V$_{[+high]}$ higher than ONSET or *V$_{[+hi]}$.ˈV and thus pronounce *camió* 'lorry' as [kə.mi.ˈo] rather than [kə.ˈmjo]). This is in line with what Recasens observes with respect to speakers who say *enlluerna* [ən^{j}.ʎu.ˈɛr.nə] 'dazzle.3SG.PRS.IND' but *enlluernador* [ən^{j}.ʎwər.nə.ˈðo] 'dazzling', with *LAPSE3σ » *COMPLEXM/V$_{[+high]}$ » ONSET.

I have not so far drawn attention to the fact that for each sequence of unstressed vocoids *both* of which are high, [ui], [iu], there are three potential syllabifications, not just two: [ui] may be [u.i], [uj], or [wi], and [iu] may be [i.u], [iw], or [ju]. In §3.2.3.2 correspondence constraints were seen preventing the second of two vocalic elements becoming a coda, and this same pattern is involved, where relevant, in the type of words exemplified in (60). *Patrioterisme* [pə.tɾju.tə.ˈɾiz.mə] 'chauvinism' contains *patriota* [pə.ˈtɾjɔ.tə] 'patriot'. To illustrate the constraints and their ranking involved in words of the type in (60) it is necessary to bring into the picture the correspondence constraint MIDCONSσ mentioned earlier, in §3.2.3.2 (27). MIDCONSσ requires the vocoid in the derivative that corresponds to [ˈɔ] in the base to be a nucleus, thereby disqualifying a candidate output where [w] corresponds to [ˈɔ]. Now recall Cabré & Prieto's claim, mentioned in §3.2.4.2, that output–output correspondence plays no role in the choice between hiatus and rising diphthong. Apart from the *cuota* pattern, this appears to be true when a high vocoid precedes a non-high one, but it is not quite right when both vocoids are high. I need to explain why a candidate like *[pətɾiwtəˈɾizmə] is rejected. I claim that MIDCONSσ does indeed give preference to a candidate where nucleus [u] in the derivative corresponds to nucleus [ˈɔ] in the base. METCONSσ (36) would achieve this effect in the variety that, indeed, has [pə.ˈtɾjɔ.tə], but it is MIDCONSσ that will prefer [pə.tɾju.tə.ˈɾiz.mə] to *[pətɾiwtəˈɾizmə] even in the 'conservative' [pə.tɾi.ˈɔ.tə] variety.

Since I am now considering possible outputs with falling diphthongs, I need to bring together some of the issues discussed earlier in §3.2.3.2. In order to make clear what is going on, and to allow us to account for alternative pronunciations where these exist, I present a set of examples and tableaux in (61). For each example I set out the variant pronunciations and the morphological relationships underlying the correspondence constraints. In each case two possibilities are considered, one with the rhythmic constraint *LAPSE3σ in the highest position, corresponding to the more informal or allegro pronunciation which I have generally been illustrating in §3.2. In the alternative variant *LAPSE3σ is inactive, representing a more formal (andante) style of pronunciation. The results with *LAPSE3σ inactive are put in parentheses.

(61) (a) *biblioteca* 'library' informal [bi.βlju.ˈtɛ.kə], formal [bi.βli.u.ˈtɛ.kə]. Output [u] corresponds to [ˈɔ] in related forms such as *bibliòfil* [bi.ˈβljɔ.fil] 'bibliophile'. (Conservative varieties with *COMPLEXM/V$_{[+high]}$ ranked above ONSET prefer [bi.βli.u.ˈtɛkə].)

biblioteca	(*LAPSE3σ)	*INCOMPM/V	METCONSσ, MIDCONSσ	*COMPM/V	ONSET	*M/V$_{[+hi]}$
(☞)bi.βli.u.ˈtɛ.kə	(*!)				*	
☞bi.βlju.ˈtɛ.kə				*		*
bi.βliw.ˈtɛ.kə			*!	*		*

(b) *campionat* [kəm.pju.ˈnat] 'championship' preferred in both informal and formal styles. Output [u] corresponds to [ˈo] in *campió* [kəm.ˈpjo] 'champion'. (Conservative varieties with *COMPLEXM/V$_{[+high]}$ ranked above ONSET prefer formal [kəm.pi.u.ˈnat] corresponding to [kəm.pi.ˈo].)

campionat	(*LAPSE3σ)	*INCOMPM/V	METCONSσ, MIDCONSσ	*COMPM/V	ONSET	*M/V$_{[+hi]}$
kəm.pi.u.ˈnat	(*!)		*!		*	
☞kəm.pju.ˈnat				*		*
kəm.piw.ˈnat			*!			*

(c) *pioner* [pi.u.ˈne] 'pioneer'. Hiatus implies a contrastively syllabic rounded vowel in the input; compare *piulet* [piw.ˈlɛt] 'chirping'.

pioner	(*LAPSE3σ)	IDENT/u/	*INCOMPM/V	*COMPM/V	ONSET	*M/V$_{[+hi]}$
☞ pi.u.ˈne					*	
pju.ˈne			*!	*		*
piw.ˈne		*!				*

(d) *arruïnar* formal [ə.ru.i.ˈna], informal [ə.ruj.ˈna], 'ruin.INF'. Output [u] and [i] correspond to the stem vowels of *ruïna* [ru.ˈi.nə] 'ruin$_N$', whose own hiatus reflects INITIALCOMPLEXM/V$_{[+high]}$.[18]

arruïnar	(*LAPSE3σ)	*INCOMPM/V	METCONSσ	*COMPM/V	ONSET	*M/V$_{[+hi]}$
(☞) ə.ru.i.ˈna	(*!)				*	
ə.rwi.ˈna			*	*!		*
(☞) ə.ruj.ˈna			*			*

(e) *almoinar* [əl.muj.ˈna] 'beg for alms'. Output [u] corresponds to [ˈɔ] in *almoina* [əl.ˈmɔj.nə] 'alms'.

almoinar	(*LAPSE3σ)	*INCOMPM/V	METCONSσ	*COMPM/V	ONSET	*M/V$_{[+hi]}$
əl.mu.i.ˈna	(*!)				*!	
əl.mwi.ˈna			*!	*		*
☞ əl.muj.ˈna						*

(f) *continuïtat* [kun.ti.nwi.ˈtat] 'continuity'. Here [w] corresponds to alternating [u] ∼ [w] in *continu* [kun.ˈti.nu] 'continuous.M' – *contínua* [kun.ˈti.nwə] 'continuous.F' with unmarked syllabification. Output [i] is part of the derivational affix /+itat/, and is subject to DERIVCONSσ (27b).

continuïtat	(*LAPSE3σ)	*INCOMPM/V	METCONSσ, DERIVCONS	*COMPM/V	ONSET	*M/V$_{[+hi]}$
kun.ti.nu.i.ˈtat	(*!)		*!		*	
☞kun.ti.nwi.ˈtat				*		*
kun.ti.nuj.ˈtat			*!			*

(g) *heroïcitat* formal [ə.ɾu.i.si.ˈtat], informal [ə.ɾuj.si.ˈtat] 'heroicness'. Output [u] corresponds to [ˈɔ] in *heroi* [ə.ˈɾɔj]; output [i] is part of the derivational affix /+ik/, and is subject to DERIVCONSσ (27b).

heroïcitat	(*LAPSE3σ)	*INCOMPM/V	METCONSσ, DERIVCONS	*COMPM/V	ONSET	*M/V$_{[+hi]}$
(☞) ə. ɾu.i.si.ˈtat	(*!)				*	
ə. ɾwi.si.ˈtat			*	*!		*
☞ ə. ɾuj.si.ˈtat			*			*

[18] I am uncertain how one would account for the normative pronunciations *amoïna* [ə.mu.ˈi.nə] 'worry.3sg.PR.IND' and *amoïnar* [ə.mu.i.ˈna], or, with *LAPSE3σ, [ə.muj.ˈna] 'worry.INF'. For *amoïna* the relevant constraints are ONSET » *COMPLEXM/V$_{[+high]}$ which would prefer [ə.ˈmwi.nə]; for

(h) *dialogar* formal [di.ə.lu.ˈɣa], informal [djə.lu.ˈɣa] 'engage-in-dialogue.INF'; cf. *dialoga* [di.ə.ˈlɔ.ɣə] 'engage-in-dialogue.3SG.PRS.IND'

dialogar	(*LAPSE3σ)	*INCOMPM/V	METCONSσ	*COMPM/V	ONSET	*M/V$_{[+hi]}$
(☞) di.ə.lu.ˈɣa	(*!)				*	
☞ djə.lu.ˈɣa		*!	*	*		*

(i) *proveïdor* formal [pɾu.βə.i.ˈðo], informal [pɾu.βəj.ˈðo] 'supplier'. Output [i] corresponds to [i] in *proveir* [pɾuβə.ˈi] 'supply.INF'.

proveïdor	(*LAPSE3σ)	*INITM/V	METCONSσ,	*COMPM/V	ONSET	*M/V$_{[+hi]}$
(☞) pɾu.βə.i.ˈðo	(*!)				*	
☞ pɾu.βəi.ˈðo			*!			*

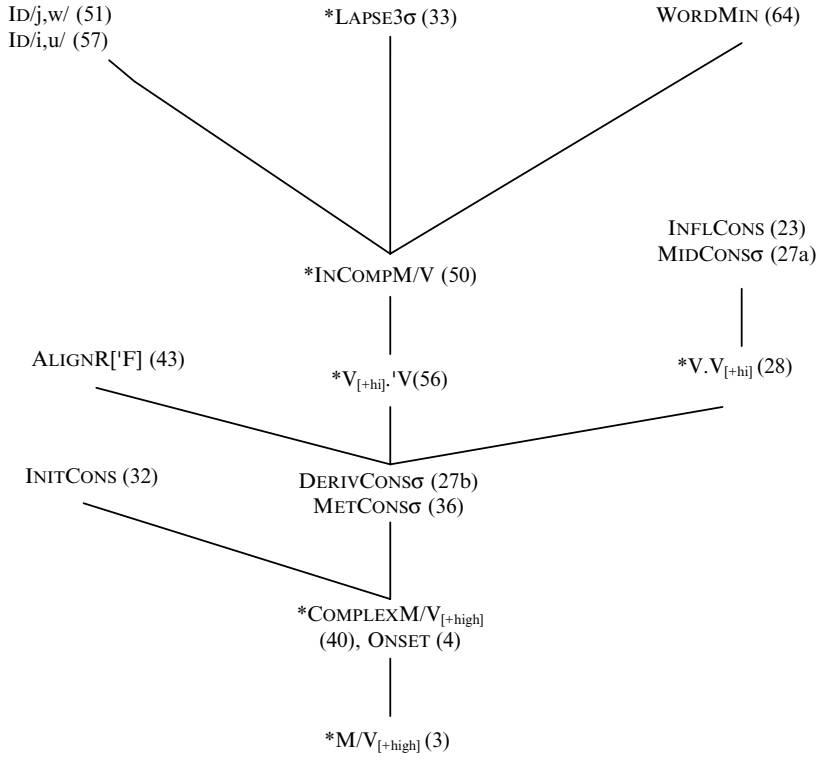
I now review the panorama of constraint interaction governing the distribution of hiatus and diphthongs. As was mentioned in §3.2.2, stressed vowels are immune to glide formation (constraint FPH: Faithfulness to prosodic head (20)), and they are not considered further at this point. I take the unmarked position, when a high vocoid is adjacent to another vocoid, to be 'no hiatus', and the basic constraint enforcing this outcome is ONSET. *COMPLEXM/V$_{[+high]}$ ranked alongside ONSET (and above *CODA) prefers a falling diphthong [ˈbuj.ðə] to a rising diphthong *[ˈbwi.ðə] or a hiatus *[ˈbu.i.ðə] (19). The metrical consistency constraint INITIAL CONSISTENCY ranking above ONSET enforces hiatus in *prohibir* [pɾu.i.ˈβi] / pro+[ib+i+R/ (31) after a prefix preceding a vocoid-initial stem. Metrical consistency of a base (METCONSσ (36)) in a derivative also overrides ONSET to impose hiatus in appropriate contexts, as in *veïnat* (35), and DERIVATIONALCONSISTENCY (27b) does likewise for e.g. *fluïdesa* (26). METCONSσ in turn is outranked by ALIGNR[ˈF] (43) to impose post-tonic rising diphthongs in *vària* (44), and by *V$_{[+hi]}$.ˈV (56) to impose a tonic rising diphthong in *variable* despite [ˈi] being a prosodic head in *varia* (57). In immediate post-tonic position DERIVATIONALCONSISTENCY (27b) is overridden by *ˈV.V$_{[+hi]}$ (28) in the case of *fluid* (29c). And *ˈV.V$_{[+hi]}$ is itself outranked by MID VOWEL CONSISTENCY (27a) to give hiatus in *etíop* (25a), and by INFLECTIONALCONSISTENCY (23) to give hiatus in *lloï*. The constraints in this first group outranking ONSET generally favour rising diphthongs in tonic and post-tonic positions, despite morphological

amoïnar METCONσ ought to preserve [i], but is irrelevant to [u], which is unstressed throughout the paradigm. Thus [ə.mwi.ˈna] should be preferred here too. Informants report that the pronunciations [ə.ˈmwi.nə] and [ə.mwi.ˈna] are, in fact, normal, thus supporting the constraints and rankings given here. I have found no other word with comparable characteristics. For the normative pronunciation one might suggest the stem [u] is lexically [+syllabic]—a similar account to what I propose for *pioner*.

relations that might favour hiatus, while morphological relations are sufficient to favour hiatus over falling diphthongs in pretonic positions.

The constraint that is the mirror image of *ˈV.V $_{[+hi]}$ (28), namely *V$_{[+hi]}$.ˈV, imposes rising diphthongs when a stressed non-high vowel is involved, in the 'general' variety (*juliol*), overriding output–output correspondence (*variable*), but is itself in turn outranked by the positional markedness constraint against initial syllable rising diphthongs *InitialComplexM/V$_{[+high]}$ (50) in, for example, *diàleg*, and this constraint prefers hiatus to rising diphthong even in unstressed initial syllables, as in *dialoga*. When the margin is not complex, as in *iuca*, or in falling diphthongs, as in *Europa*, it is Onset that penalizes hiatus. Faithfulness to underlying /w/ in lexical 'labiovelar' clusters (37) overrides the hiatus-promoting constraints, *InitialComplexM/V$_{[+high]}$ and more generally *ComplexM/V$_{[+high]}$, and it is likewise faithfulness to underlying glides that accommodates lexical exceptions like those in (37) (and rare complex pretonic codas as in *Austràlia* [əws.ˈtɾa.ljə] 'Australia'. The rhythmic constraint *Lapse3σ, which is partly governed by style or speech rate, tends to favour diphthongs, both falling and rising, over hiatuses when the alternative is a string of three or more unstressed syllables. So the overall pattern of constraint ranking is as given in (62).

(62)



3.2.5 *High vocoid sequences*

Taking up an issue that was a major focus of §3.2.4.8, one can ask what generally governs the syllabification of sequences of two adjacent vocoids both of which are high.

3.2.5.1 *Adjacent dissimilar high vocoids*

Consider first cases where the two adjacent high vocoids are different, that is /iu/ or /ui/. In part this ground has been covered in §3.2.4.4, where the example *iuca* 'yucca' was given (51), illustrating the ranking ONSET, *COMPLEXM/$V_{[+high]}$ » *M/$V_{[+high]}$. A few more examples (63) illustrate the same syllabification in initial position:

(63) (a) *iugoslau* [ju.ɣuz.ˈlaw] 'Yugoslav'
iogurt [ju.ˈɣurt] 'yogurt'
whisky [ˈwis.ki] 'whisky'
UIB [ˈwip] (acronym of *Universitat de les Illes Balears*)
(b) *Iu* [ˈiw] (given name, = Yves)
ui [ˈuj] 'ow!'
oïda [u.ˈi.ðə] 'hearing'

The question now is, why are the examples in (63b) not *[ˈju], *[ˈwi], and *[ˈwi.ðə] respectively? For *Iu* and *ui* a word-minimality condition (64) seems plausible.

(64) WORDMIN: A lexical word contains no fewer than two moras.[19]

A nucleus plus coda [iw] or [uj] conforms to this constraint as an onset plus nucleus does not. WORDMIN also serves to account for the hiatus in *ió* [i.ˈo] 'ion', which is one of the few words generally agreed to have an initial syllabic [i] in hiatus. As function words, *jo* [jɔ] 'I' and *ja* [ja] are not subject to this constraint (nor are clitic pronouns such as *lo* [lu] '3SG.M.OBJ', *ne* [nə] 'partitive'). A few other words have initial hiatus in the variety described by Cabré & Prieto, such as *IEC* [i.ˈɛk] (acronym of *Institut d'Estudis Catalans*) and *hiat* [i.ˈat] 'hiatus' (which for other speakers is [ˈjat]). But no markedness constraint preferring [i.ˈɛk] and [i.ˈat] could distinguish these from *iac* [ˈjak] 'yak', *ien* [ˈjɛn] 'yen', *iot* [ˈjɔt] 'yacht', and so on. There is no alternative, I think, to acknowledging a lexical contrast between /i/ and /j/ in initial prevocalic position (with /i/ being the marked member), a contrast of the type occasionally seen in this chapter. Likewise there is no alternative to lexically marked contrast to account for /w/ vs. /u/ in the examples of (65).

(65) *UEFA* [ˈwɛ.fə] 'UEFA' *oasi* [u.ˈa.zi] 'oasis'
western [ˈwɛs.tərn] 'western (film)' *oest* [u.ˈest] 'west'

[19] WORDMIN may well reduce to LEXWD = PROSODICWORD together with FOOTBINARITY.

3.2.5.2 *Adjacent similar high vocoids*

What is the position with regard to syllabification when two adjacent high vocoids have the *same* place features: /ii/ or /uu/? I assume, perhaps uncontroversially, that, within one syllable, identical elements cannot occur more than once in the same syllable position. Thus Onset+Onset [$_{\sigma}$uu] is excluded (—the least violable aspect of OCP?). So the options to be considered are those set out in (66).

(66) (a) Nucleus+Nucleus (two syllables), symbolized [i.i], [u.u]
(b) Onset+Nucleus, symbolized [ji], [wu]
(c) Nucleus+Coda, symbolized [ij], [uw] (equivalently long vowels [iː], [uː])
(d) Nucleus+Onset, symbolized [i.j], [u.w]
(e) Coda+Onset, symbolized [j.j], [w.w] (geminate glides)

Though each of the patterns (66a–d) can be justified in Catalan words (or clitic groups) for morphological reasons (and may thus be distinguished in the mental representations of speakers), possibly only (66a), the bisyllabic sequence, can be distinguished by listeners. (Geminate glides (66e) are not found within morphemes, and there are no affixation patterns that would give rise to them. They may arise in sandhi, of course; see Chapter 4.)

The Nucleus+Nucleus pattern (66a) can be morphologically justified in items such as those in (67).

(67) *xiïta* [ʃi.ˈi.tə] /ʃi+ˈitə/ 'Shiite'
nihilisme [ni.i.ˈliz.mə] 'nihilism'
fiï [ˈfi.i] /ˈfi+i/ 'trust.3SG.PRS.SUBJ'
oologia [u.u.lu.ˈʒi.ə] /uˈɔ+ˈlɔʒ+ˈi+ə/ 'oology'
actuo [ək.ˈtu.u] /əkˈtu+u/ 'act.1SG.PRS.IND'
duodè [du.u.ˈðɛ] 'duodenum'
alcoholisme [əl.ku.u.ˈliz.mə] /əlkuˈɔ l+ˈizmə/ 'alcoholism'
fluorescent [flu.u.ɾə.ˈsen] (also [uw], [u] Recasens 1993: 107, 110) /fluˈor+əs+ˈent/ 'fluorescent'
cooperar [ku.u.pə.ˈɾa] (also [u]) /ku+uˈpeR+ˈa+R/ 'cooperate'
afectuositat [ə.fək.tu.u.zi.ˈtat] (also [u] Recasens 1993: 103) /əˈfekt+u+ˈoz+iˈtat/ 'affectionateness'

In such cases the two syllabic vocoids can be justified either as each belonging to different morphemes (and preserved from glide formation by INFLECTIONAL CONSISTENCY (23)—as in *actuo* or *fiï*—or by METCONSσ, as in *afectuositat*, etc.); or by the fact that one corresponds to a non-high head in the base by MIDCONSσ (*alcoholisme*, etc.); or in the exceptional cases (such as *nihilisme*, *duodè*) by lexical marking of syllabicity. The versions with reduction of [u.u] to [u] are entirely in line with the effect of the *LAPSE3σ constraint, reducing sequences of more than two unstressed syllables. I believe there is no audible difference, for Catalan, between [uw] and [u], given undominated $^{*}V_{\mu\mu}$.

The Onset+Nucleus sequence (66b) can be illustrated by examples such as those in (68). A high vocoid is followed by a vocoid, in a context which would generally allow glide formation, to satisfy ONSET.

(68) *desmaï* [dəz.ˈma.i] 'faint.3SG.PRS.SUBJ' /dezˈmai+i/ (the orthography *desmaï* rather than **desmaii* reflects the realization of /Vii/ as [V.i])
remeï [rə.ˈmɛ.i] 'make-good.3SG.PRS.SUBJ' /reˈmɛi+i/
heroïna [ə.ɾu.ˈi.nə] 'heroine' /eˈɾɔi+ˈin+ə/
hi inicià [i.ni.ˈsja] 'began.3SG.PST.PRF there' (Recasens 1993: 134)
pouo ∼ *poo* [ˈpo.u] 'draw (water).1SG.PRS.IND' /ˈpou+u/
creuo [ˈkɾɛ.wu] ∼ [ˈkrɛ.u] 'cross.1SG.PRS.IND'
adequo [ə.ˈðɛ.kwu] ∼ [ə.ˈðɛ.ku] 'accommodate.1SG.PRS.IND'
enaiguo [ə.ˈnaj.ɣwu] ∼ [ə.ˈnaj.ɣu] 'water.1SG.PRS.IND'
quotidià [kwu.ti.ˈðja] ∼ [ku.ti.ˈðja] (also [kwo-]) 'everyday'
escriu-ho [əs.ˈkɾi.wu] ∼ [əs.ˈkɾi.u] 'write.2SG.IMP it' (Recasens 1993: 140)
ho usava [u.ˈza.βə] 'use.3SG.PST.IMP it' (Recasens 1993: 134)

What the alternative transcriptions reveal in the examples of (68) is perhaps, above all, the difficulty of satisfying competing requirements of morphological and phonetic perceptual distinctiveness. Such conflicts are confronted by an approach such as that of Boersma (2000), who argues for distinct production and perception grammars. To illustrate the issues, take the case of *escriu-ho* 'write.2SG.IMP it'. The speaker intends to concatenate two morphemes, which for the sake of argument I represent here as /ˈskriu/ and /u/. The first of these is a word in its own right, with a regular pronunciation [əs.ˈkɾiw] when it occurs in isolation or followed by a consonant-initial word, and essentially the same pronunciation (even if syllabified differently, for example, in *escriu el teu nom* [əs.ˈkɾi.wəl.tew.ˈnɔm] 'write your name') when followed by a vowel-initial word. The clitic /u/ [u] ∼ [w] always gets syllabified according to its phonological context. So for the sequence, the speaker intends to utter [əs.ˈkri.wu], concatenating the regular forms of the two morphemes in a sequence that respects ONSET. The problem is that there are no features of phonetic production that distinguish Catalan [w] from [u].[20] So the phonetic input cannot be distinguished from [əs.ˈkri.u] (which contains a perceptual ONSET violation). From this perspective, the conventional orthographic representation of verb forms such as *desmaï* [dəz.ˈma.i] 'faint.3SG.PRS.SUBJ' /dəzˈmai+i/ institutionalizes the output of the 'perception grammar'. To some extent, no doubt, this is because the schema ˈCV.i is well established in the present subjunctive of conjugation I verbs with vowel final stems (such as *creï* [kɾe.i] /kɾe+i/ 'create3SG.PRS.SUBJ'). It is also the case that what might be represented as [ji] does not arise at all

[20] The situation in English is quite different, where /iː/ and /uː/ are articulatorily and perceptually rather distinct from /j/ and /w/ respectively—the vowels being lower, more centralized, and /uː/ less rounded. So in English 'yeast' /jiːst/ and 'woo' /ˈwuː/ are quite adequately distinguishable from 'east' /iːst/ and 'ooh' /uː/.

frequently through morphological concatenation, though there are a few other examples like *heroïna* in (68).

The Nucleus+Coda sequence (66c) and Nucleus+Onset sequence (66d) can be considered together. The representation of Nucleus+Coda [ij], [uw], is equivalent to that of long vowels [iː], [uː]. In autosegmental terms, a set of vowel place features is linked to the two moras of a branching rhyme. And Catalan has no vowel length contrast, except as the consequence of a sequence of two nuclei, as mentioned above (66a). So 'morphological' [u+w] is produced, and perceived, as [u] (69a). In [u.wV], where the onset is not moraic, again there is no perceptual difference from [u.V]. Examples in (69) illustrate this.

(69) (a) *duu* ∼ *du* [ˈdu] 'bring.3SG.PRS.IND' /ˈduu/
duen [ˈdu.ən] 'bring.3PL.PRS.IND' /ˈduu+ən/
mourà [mu.ˈɾa] 'move.3SG.FUT' /ˈmɔu+R+ˈa/ with vowel reduction of unstressed /ɔ/ to [u]
plouria [plu.ˈɾi.ə] 'rain.3SG.COND' /ˈplɔu+R+ˈiə/
(b) *pouar* ∼ *poar* [pu.ˈa] 'draw (water).INF' /ˈpou+ˈa+R/
bouada [bu.ˈa.ðə] 'herd of oxen' /ˈbɔu+ˈad+ə/
ouet [u.ˈɛt] 'egg.DIM' /ˈɔu+ˈɛt/

(The absence of contrasting vowel length shows that the constraint $^{*}V_{\mu\mu}$ is active in Catalan.) There are no word-internal examples where the sequence [ij] or [i.j] might be justified morphologically. Recasens (1993: 133) gives a transcription [ij] for *vagi-hi* [ˈba.ʒij] 'go.3SG.IMP-there'. The precise phonetic realization is unclear to me.

4

PHRASAL PHONOLOGY: VOWEL SANDHI

4.1 INTRODUCTION

The issue of vowel sandhi in Catalan is one of considerable complexity and challenge. Chapter 4 in Recasens (1993) is a rich (and dense) source of data, and of interpretation; the data I treat here is largely from this source. To see the wood for the trees is no small challenge. Many words in Catalan are vowel-initial or vowel-final. When a vowel-final and vowel-initial word are adjacent in a phrase, one of four possible outcomes is at issue:

- hiatus (no change to the shape of the words as uttered in isolation),

and three types of syllable reductions which are deviations from 'faithfulness':

- glide formation (if one of the vowels in contact is high),
- fusion (if the vowels in contact are identical), and
- elision. (Fusion might be thought to be subsumed under elision, but for the moment it is convenient to keep them apart.)

Broadly speaking we might expect unfaithfulness to a V.V sequence, involving reduction, to be driven by the ONSET constraint (syllables should have onsets: see Chapter 3). In fact, ONSET hardly ever proves to be the crucial constraint. Rather, syllable reduction is driven

(a) by members of the *LAPSE constraint family, which penalize unstressed prosodic material (syllables, moras) between stressed nuclei.
(b) by constraints of the OCP (Obligatory Contour Principle) family which penalize immediate repetition of identical phonological material.
(c) by, in some cases, constraints penalizing Və sequences, where the issue is that the least prominent vowel [ə] is masked by the qualities of the preceding vowel, often a stressed one, while [ə] is always unstressed.

By contrast, we might expect hiatus to be maintained

(a) by Correspondence constraints (penalizing deviations from inputs or from the forms of words in 'basic' contexts). Correspondence constraints do have an important role, a fact that was recognized by Palmada (1994a: 125), though expressed then in rather different pre-OT terms.

(b) by constraints of the *CLASH family, which penalize sequences of two stressed syllables without sufficient unstressed prosodic material (syllables, moras) between them. (For Catalan, the ideal rhythmic pattern appears to consist of stresses separated by at least one mora and not more than three.)
(c) by Faithfulness to Prosodic Heads (FPH): stressed vowels are not subject to significant modification, as in *vaig demanar ànec* [ˈbadʒ.ðə. mə.ˈna.ˈa.nək] 'I asked for duck'. Little more will be said about contacts between stressed vowels here, though Recasens (1993: 124–6) gives information about the possible fusion or synaeresis of stressed vowels in certain contexts.

The major task of this chapter is to show how members of the above constraint families, whether promoting or restraining syllable reduction, are interleaved.

4.2 STRESSED VOWEL FOLLOWED BY UNSTRESSED VOWEL

4.2.1 *Stressed vowel followed by an unstressed high vowel*

Consider first cases where a stressed vowel is followed by an unstressed high vowel of different quality as in (1). The unstressed high vowel is realized as a glide in (1b) and (1c) but not in (1a).

(1) (a) *menú idoni* [mə.ˈnu.i.ˈdɔ.ni] 'suitable menu'
proferí udols [pɾu.fə.ˈɾi.u.ˈðols] 'utter.3SG.PST.PRF howls'
(b) *menú imprès* [mə.ˈnujm.ˈpɾɛs] 'printed menu'
això importa [ə.ˈʃɔjm.ˈpɔr.tə] 'that matters'
(c) *destí opcional* [dəs.ˈtiwp.sju.ˈnal] 'optional fate'
tabú inaudit [tə.ˈβuj. nəw.ˈðit] 'outrageous taboo'

The differing syllabifications of the examples in (1) illustrate several considerations that recur in the rest of the chapter. The examples in (1a) show hiatus (violating ONSET) preferred to *CLASH violations.

(2) *CLASH: Phonological phrase-head stresses do not fall on adjacent syllables whose heads are separated by no more than one mora.

The *CLASH constraint is formulated here in terms of phonological phrases because the relevant cases are ones that involve two contiguous phonological phrases with lexical heads, NP AP, NP VP, and so on. (The point of the subordinate clause in the definition will be addressed shortly.) *CLASH does not penalize diphthongization or glide formation when the first word is in pre-lexical head position, as in *va ofendre* [ˈbaw.ˈfɛn.dɾə] 'AUX.3SG.PST.PRF offend', *no oblidis* [ˈnow.ˈβli.ðis] 'not forget.2SG.SUBJ'. *darrer udol* [də.ˈrew.ˈðol] 'last howl'. In these sequences there is only one phonological phrase. Likewise, a 'restructured' phonological phrase, involving, say, a verb of functional import and its complement, easily tolerates adjacent stresses, as in *té idees* [ˈtej.ˈðe.əs] 'has ideas', *fa ometre* [ˈfaw.ˈmɛ.trə] 'causes to omit'.

In (1a) the high initial unstressed vocoids are in open syllables: hiatus results. In (1b), however, the high initial unstressed vocoids are in closed (heavy) syllables, and glide formation takes place. Stress on adjacent syllables is tolerated here; in fact, it is preferred to hiatus. More precisely, a complex coda, with presumably two moras, provides sufficient distance between the syllable nuclei of adjacent stressed syllables to allow no clash to be counted. The (1c) examples also have unstressed initial vowels in open (light) syllables, like (1a), but here *CLASH is not at issue, and glide formation takes place.

A similar pattern to (1) is found when the adjacent high vowels are of the same quality. Here, though, diphthong formation is not an option (see §3.2.5.2; nor is vowel length an option, Catalan having a high ranking constraint $*V_{\mu\mu}$). Elision/fusion is the only viable alternative to hiatus. Examples in (3a–c) match those in (1) in relevant respects, here with elision/fusion in (3b) and (3c).

(3) (a) *destí idoni* [dəs.ˈti .i.ˈðɔ.ni] 'suitable fate'
comú oblit [ku.ˈmu .u.ˈβlit] 'common oblivion'
(b) *destí incert* [dəs.ˈtin.ˈsɛrt] 'uncertain fate'
tabú obscur [tə.ˈβups.ˈkur] 'obscure taboo'
(c) *tabú oblidat* [tə.ˈβu.βli.ˈðat] 'forgotten taboo'
submarí inactiu [ˈsum.mə.ˈɾi.nəkˈtiw] 'inactive submarine'

What is it that allows elision in heavy syllables (3b), despite the *CLASH violations? We may recall the constraint mentioned in §3.2.3.2 (28), which penalizes hiatus between a stressed vowel and a high vowel within words ($*^{\prime}V.V_{[+high]}$), but the (1a) examples show that this constraint can be overridden, provided the vowels are of different quality, and in separate words.

What seems to be involved additionally in (3b) is an OCP-type constraint (4) penalizing hiatus between vowels of the same quality ($*V_{\alpha}.V_{\alpha}$). When, as here, the first vowel is stressed, such a constraint would have a clear perceptual motivation, in that a separate, less prominent vowel of the same quality as the preceding vowel is hard to distinguish (just as, in the case of $*^{\prime}V.V_{[+high]}$ in §3.2.3.2, it was argued that the syllabic nature of the second vowel is hard to distinguish in this context.)

(4) $*OCP/^{\prime}V_{\alpha}.V_{\alpha}$: Two vowels with the same place features, the first being stressed, do not stand adjacent to one another.

Examples like *destí incert* (3b) demonstrate the ranking $*OCP/^{\prime}V_{\alpha}.V_{\alpha}$ » *CLASH. But the examples like *comú oblit* (3a) show in turn that $*OCP/^{\prime}V_{\alpha}.V_{\alpha}$ must be subordinate to another constraint. Recall that the second elements of (3a) begin with open syllables. Elision/fusion here would produce a 'worse' clash than that found in types (3b) or (1b), namely, one with stress on adjacent moras, not simply on syllables whose heads are separated by one mora. In (1b) the head moras of adjacent stressed syllables are separated by two moras with a complex coda resulting from glide formation: no clash is counted. In (3b) the stressed syllable resulting from elision/fusion is heavy; the head mora of the syllable is

separated from the head mora of the next, stressed syllable by a coda mora. In (3a), however, if elision/fusion took place (glide formation/lengthening being ruled out by *$V_{\mu\mu}$), stress would fall, fatally, on *adjacent* moras: *destí idoni* *[dəs.ˈti.ˈðɔ.ni] ≺ [dəs.ˈti.i.ˈðɔ.ni], *comú oblit* *[ku.ˈmu.ˈβlit] ≺ [ku.ˈmu.u.ˈβlit]. What this shows is the high ranking relative to other constraints considered so far in this section of the *CLASHμ constraint (5)—inherently outranking the more general *CLASH. That is to say, the constraint ranking is *CLASHμ » *OCP/ˈV_α.V_α » *CLASH.

(5) *CLASHμ: Phonological phrase-head stresses do not fall on adjacent moras.

In tableau (6) evaluation of the candidates for *menú idoni* (1a) shows the ranking *CLASH » ONSET. The evaluation for *menú imprès* (1b) shows the ranking ONSET over an output–output correspondence constraint such as UNIFORMσ (§3.1.1 (15)). The example *destí incert* (3b) shows *OCP/ˈV_α.V_α » *CLASH; and in (6d) *comú oblit* (3a) shows *CLASHμ » *OCP/ˈV_α.V_α.

(6)

		*CLASHμ	*OCP/ˈV_α.V_α	*CLASH	ONSET	UNIσ
(a)	*menú idoni*					
	☞ mə.ˈnu .i.ˈðɔ.ni				*	
	mə.ˈnuj.ˈðɔ.ni			*!		*
(b)	*menú imprès*					
	mə.ˈnu .im.ˈpɾɛs				*!	
	☞ mə .ˈnujm.ˈpɾɛs					*
(c)	*destí incert*					
	dəs.ˈti .in.ˈsɛrt		*!		*	
	☞ dəs.ˈtin.ˈsɛrt			*		*
(d)	*comú oblit*					
	☞ ku.ˈmu .u.ˈβlit		*		*	
	ku.ˈmu .ˈβlit	*!		*		*

*CLASHμ » *OCP/ˈV_α.V_α » *CLASH » ONSET » UNIFORMσ

Among the examples in this chapter so far, FPH output–output correspondence issues have not come up: none of the syllables subject to glide formation or elision/fusion has corresponded to a stressed syllable in a base. Recasens' data (1993: 126–8) show that initial stress in a base can in fact block glide formation or fusion/elision provided the resulting distance between stressed syllables is not more than two moras, as in (7a) and (7b). (Example (3c), *tabú oblidat* [tə.ˈβu.βli.ˈðat] ≻ *[tə.ˈβu.u.βli.ˈðat], has already shown that a two-mora gap with hiatus is not tolerated when there is no support from a stressed vowel

in a base.) Examples (7c) and (7d) show that a three-mora gap involving hiatus is disfavoured, whatever the morphological structure of the second word.

(7) (a) *hindú honrat* [in.ˈdu .$u_{\mu}n_{\mu}$.ˈrat] 'honest Hindu', cf. *honra* [ˈon.rə] 'honour'
(b) *hindú humanista* [in.ˈdu .u_{μ}.mə$_{\mu}$.ˈnis.tə] 'humanist Hindu', cf. *home* [ˈɔmə] 'man'
(c) *hindú honradíssim* [in.ˈdun.rə$_{\mu}$.ˈði.sim] ≻ *[in.ˈdu. $u_{\mu}n_{\mu}$.rə$_{\mu}$.ˈði.sim] 'most honest Hindu', cf. *honra* [ˈon.rə] 'honour'
(d) *destí unificat* [dəs.ˈtiw.ni$_{\mu}$.fi$_{\mu}$.ˈkat] ≻ *[dəs.ˈti .u_{μ}.ni$_{\mu}$.fi$_{\mu}$.ˈkat] 'unified fate', cf. [ˈun] 'one'

The correspondence constraint enforcing syllabicity of the initial vowel in cases like (7a) and (7b) is similar to the one mentioned previously (§3.2.3.3: (36)) enforcing syllabicity in a derived word. A different version is given here to take into account a different correspondence relation, namely, between base and phonological phrase.

(8) METRICAL CONSISTENCYBase-PPhrase (METCONSB-Phr): The head of a head foot in a base corresponds to the head of a syllable in a phonological phrase. (An output–output correspondence version of Faithfulness to Prosodic Heads (FPH).)

Comparison of (7a) and (7b) with (3a) and (3b) demonstrates the ranking METCONSB-Phr » *OCP/ˈV_{α}.V_{α}. Comparison of (7a) and (7b) with (7c) and (7d) shows that the *LAPSE constraint *LAPSE3μ (9)[1] dominates METCONSB-Phr. Actually, the activity of rhythm constraints like *LAPSE3μ depends on speech rate. Their effects may be suppressed with a more formal or deliberate speech rate.

(9) *LAPSE3μ: Between the head syllables of prosodic words three moras do not intervene.

Before moving on from sequences of stressed high vowel followed by unstressed high vowel, I should mention a further consideration favouring hiatus: when the stressed vowel is that of a monosyllabic word belonging to a major lexical category (noun, adjective, or verb). The examples in (11) resemble prosodically those of (1b), *menú imprés* [mə.ˈnujm.ˈprɛs] 'printed menu', and (1c), *destí opcional* [dəs.ˈtiwp.sju.ˈnal] 'optional fate', but show hiatus rather than fusion or glide formation.

[1] Or, high-ranking *LAPSE3σ in (7d).

(10)

		*LAPSE3μ	METCONSB-Phr	*OCP/'V_α.V_α	*CLASH	ONSET
(7a)	*hindú honrat* cf. *honra* [ˈon.rə]					
	☞ in.ˈdu .$u_\mu n_\mu$.ˈrat			*		*
	in.ˈdun.ˈrat[2]		*!		*	
(7b)	*hindú humanista* cf. *home* [ˈɔ.mə]					
	☞ in.ˈdu. u_μ.$mə_\mu$.ˈnis.tə			*		*
	in.ˈdu.mə .ˈnis.tə		*!			
(7c)	*hindú honradíssim* cf. *honra* [ˈon.rə]					
	in.ˈdu. $u_\mu n_\mu$.$rə_\mu$.ˈðɪ.sim	*!		*		*
	☞ in.ˈdun.$rə_\mu$.ˈð ɪ.sim		*			

*LAPSE3μ » METCONSB-Phr » *OCP/'V_α.V_α

(11) *va dur infants* [ˈba .ˈðu .iɲ.ˈfans] 'AUX.3SG.PST.PRF take children'
vi importat [ˈbi .ˈim.pur.ˈtat] 'imported wine'
dur ultratge [ˈdu .ul.ˈtrad͡ʒə] 'serious insult'

These cases will be considered alongside similar examples with an initial non-high vowel in the contact position.

4.2.2 *Stressed vowel followed by an unstressed non-high vowel*

In eastern continental Catalan, which is primarily considered here, as a result of vowel reduction in unstressed syllables the only unstressed non-high vowel in initial position is [ə]. The possibilities in the case of input ˈV##ə are hiatus or elision. When the lowest, most prominent vowel, [ˈa] is followed by a word beginning with [ə], elision is obligatory (12a). With the other low vowels, [ˈɛ] and [ˈɔ], marginally less prominent than [ˈa], elision also seems to be practically obligatory (12b).

(12) (a) *està atordit* [əs.ˈta .tuɾ.ˈðit] 'is stunned'
mesurar alçària [mə.zu.ˈɾal.ˈsa.ɾjə] 'measure.INF height'
un faisà enorme [uɲ.fəj.ˈza.ˈnor.mə] 'a huge pheasant'

[2] The evaluation of this candidate assumes that the two input [u]s are not fused into the one [u] of the output. With fusion, METCONS would score no violation. I infer here that UNIFORMITY dominates METCONS.

(b) *cafè amarg* [kə.ˈfɛ.ˈmark] 'bitter coffee'
això anima [ə.ˈʃɔ.ˈni.mə] 'that encourages'

The relative prominence of Catalan vowels has been discussed in §2.3.7. What appears problematic about the realization of [ˈa.ə], [ˈɛ.ə], and [ˈɔ.ə] is that a stressed vowel of high perceptual prominence is followed by the unstressed vowel of minimum prominence, [ə]. Perceptually, it is as if the prominent vowel overwhelms the non-prominent vowel, or obscures its presence entirely. Because such sequences are hard to perceive, speakers may not take the trouble to make the sequence at all, despite loss of semantic information in some cases. Or, following the perspective mentioned in §3.2.5.2, one might alternatively suggest that [ˈa.ə] etc. are grammatical for the production grammar, but that such sequences cannot be reliably identified by the perception grammar. The relevant constraint is simply abbreviated here, as *ˈV_{Low}.ə (13). It is a specific version of *ˈV.ə, which it inherently outranks and to which I return.

(13) *ˈV_{Low}.ə: *ɛ.ə, *ɔ.ə, *a.ə

The constraint *ˈV_{Low}.ə, ranking above *CLASHμ, rules out hiatus in examples such as those of (12). Within words such [ˈɛ.ə] and [ˈɔ.ə] sequences are attested (e.g. *proa* [ˈpɾɔə] 'prow'), showing that *ˈV_{Low}.ə is subordinate to faithfulness constraints such as CONTIGUITY, and MAXI-OFeatures, or MAXMorpheme.

Other vowel sequences involving word-initial [ə] illustrate patterns of elision and hiatus that match those seen in §4.2.1. Matching (1a) *menú idoni*, the examples in (14) reflect the ranking *CLASHμ » ONSET. Hiatus after a stressed high or mid vowel is better than a *CLASHμ violation.

(14) *tabú atàvic* [tə.ˈβu .ə.ˈta.βik] 'atavistic taboo'
actor enèrgic [ək.ˈto .ə.ˈnɛr.ʒik] 'energetic actor'
escriptor agut [əs.kɾip.ˈto .ə.ˈɣut] 'penetrating writer'
destí amarg [dəs.ˈti .ə.ˈmark] 'bitter fate'
no animes [ˈno .ə.ˈni.məs] 'not encourage.2SG.PRS.IND'.

(15) *actor esplèndid* [ək.ˈtos.ˈplɛn.dit] 'splendid actor'
jardí antic [ʒəɾ.ˈðin.ˈtik] 'old garden'
collir espàrrecs [ku.ˈʎis.ˈpa.rəks] 'pick.INF asparagus'
cosí astròleg [ku.ˈzis.ˈtrɔ.lək] 'astrologist cousin' (ignoring base form *astre* 'heavenly body'—with *astre* in mind [ku.ˈzi .əs.ˈtrɔ.lək] is possible).

Examples in (15) match (1b) *menú imprès*. Hiatus involving [ə] after a high or mid vowel is worse than a *CLASH violation. It is in these examples that the effect of the more general constraint against schwa preceded by a stressed vowel comes into evidence. This general version of *ˈV_{Low}.ə is spelt out in (16).

(16) *ˈV.ə: No stressed vowel followed by schwa.

Examples in (17a), (17b) correspond to those in (7a) (*hindú honrat*), (7b) (*hindú humanista*) respectively. They demonstrate the ranking METCONSB-Phr » *ˈV.ə.

Example (17c) matches (7c) (*hindú honradíssim*) in showing *LAPSE3μ » MET-CONSB-Phr. The examples in (17d), which match (1c) (*destí opcional*), illustrate evaluation by *LAPSE3μ. Those of (17e), which match (3c) (*tabú oblidat*) in relevant respects, show the effect of *ˈV.ə » ONSET, unrestrained by correspondence considerations. They contrast with the (17a), (17b), and (17c) cases.

(17) (a) *corder anyal* [kuɾ.ˈðe .ə.ˈɲal] 'lamb less than one year old', cf. *any* 'year'
setí arnat [sə.ˈti .ər.ˈnat] 'moth-eaten satin', cf. *arna* 'moth'
cançó artística [kən.ˈso .ər.ˈtis.ti.kə] 'artistic song', cf. *art* 'art'
(b) *sentir avidesa* [sən.ˈti .ə.βi.ˈðɛ.zə] 'feel.INF greed', cf. *àvid* 'greedy'
(c) *actor alpinista* [ək.ˈtol.pi.ˈnis.tə] 'mountaineering actor', cf. *Alp* 'Alp'
(d) *tabú acceptable* [tə.ˈβuk.səp.ˈtab.blə] 'acceptable taboo'
tauló allargat [təw.ˈlo.ʎər.ˈɣat] ≻ *[təw.ˈlo .ə.ʎər.ˈɣat] 'long plank'
(e) *camí agradable* [kə.ˈmi.ɣɾə.ˈðab.blə] 'pleasant path'[3]
pastor alemany [pəs.ˈto.lə.ˈmaɲ] 'German shepherd'

I return to consider the case of hiatus after major lexical monosyllables illustrated in (11) above and repeated here in (18a) and (18b), alongside parallel cases with initial [ə] in the second word.

(18) (a) *va dur infants* [ˈba .ˈðu .iɱ.ˈfans] 'AUX.3SG.PST.PRF take children'
dur ultratge [ˈdu .ul.ˈtɾad͡ʒə] 'serious insult'
fi estranya [ˈfi .əs.ˈtra.ɲə] 'unusual end'
sí emfàtic [ˈsi .əɱ.ˈfa.tik] 'emphatic yes'
(b) *vi econòmic* [ˈbi .ə.ku.ˈnɔ.mik] 'inexpensive wine'
té animals [ˈte .ə.ni.ˈmals] 'has animals'
fi agradable [ˈfi .ə.ɣɾə.ˈðab.blə] 'pleasant end'
(c) *vi importat* [ˈbi .im.pur.ˈtat] 'imported wine'
vi exquisit [ˈbi .əks.ki.ˈzit] 'select wine'
(d) *fi astoradora* [ˈfis.tu.ɾə.ˈðo.ɾə] 'appalling end'
vi extraordinari [ˈbiks.tɾəwɾ.ði.ˈna.ɾi] 'extraordinary wine'

The examples in (18a) resemble those in (3b) in prosodic and morphological structure. However, the prevocalic stressed vowel is in a major-category monosyllable (as opposed, for example, to a pronoun such as occurs, with elision, in *tu escrius* [ˈtus.ˈkɾiws] 'you.SG write.2SG.PRS.IND'). The examples in (18b) are parallel to those in (17e). In the (18a) and (18b) types hiatus is the norm, as it is also in (18c). A possible interpretation of the relevant constraint is that the overlapping of two word-initial syllables is an especially difficult challenge for word recognition. We have seen many examples in this section which show that preservation of individual initial syllables is not of crucial importance, but in the examples previously considered, an initial syllable was lost after, or merged with, a non-initial syllable of a previous word. The requirement here seems to be that adjacent *initial* syllables of major category words (or perhaps adjacent initial

[3] For Recasens the prosodically parallel *matí avorrit* 'boring morning' has hiatus.

syllables of minimal phonological phrases) should remain adjacent and not overlap. There are no doubt several ways in which this constraint might be formulated. Pending further investigation, the following statement is offered provisionally (19). It is an anchoring constraint.

(19) *InitialOverlap (*InOv): The right edge of a lexical word which consists of a degenerate foot (ˈL) corresponds to the right edge of a syllable.

Examples (18a) and (18b) show that *InitialOverlap ranks above *OCP/ ˈV_α.V_α. The examples *vi importat*, *vi exquisit* (18c) show that *InitialOverlap ranks higher than *Lapse3μ (9). Elision in examples (18d), however, shows that *InitialOverlap is itself dominated by the yet more specific *Lapse constraint *Lapse3σ (§3.2.3.3 (33)). That is to say, the constraint ranking demonstrated by the set of examples in (18) is as illustrated in tableau (20).

(20)

		*Lapse3σ	*InOv	*Lapse3μ	*ˈV.ə	*Clash	Onset
(a)	*fi estranya*						
	☞ ˈfi .əs.ˈtra.ɲə				*		*
	ˈfis.ˈtra.ɲə		*!			*	
(b)	*fi agradable*						
	☞ ˈfi .ə. ɤɾə.ˈðab.blə				*		*
	ˈfi.ɤɾə.ˈðab.blə		*!				
(c)	*vi exquisit*						
	☞ ˈbi .əks.ki.ˈzit			*	*		*
	ˈbiks.ki.ˈzit		*!				
(d)	*fi astoradora*						
	ˈfi .əs.tu.ɾə.ˈðo.ɾə	*!		*	*		*
	☞ fis.tu.ɾə.ˈðo.ɾə		*				

*Lapse3σ » *InitialOverlap » *Lapse3μ » *ˈV.ə

This completes the survey of word sequences involving a stressed vowel followed by an unstressed vowel. The next section, §4.3, turns to sequences involving an unstressed vowel followed by a stressed vowel.

4.3 UNSTRESSED VOWEL FOLLOWED BY STRESSED VOWEL

When an unstressed vowel that is high ([i] or [u]) precedes a stressed vowel, the potential outcomes are: fusion before a high vowel of the same quality, a rising

diphthong before any other vowel quality, or, of course, in either case, hiatus. Except in absolute phrase-initial position (21) where we see glide formation, hiatus seems to be the standard outcome.

(21) *ho obre* [ˈwɔ.βɾə] 'opens it'
i ara m'ho diu [ˈja.ɾə.mu.ˈðiw] 'and now tell.3SG.PRS.IND me it'
hi entres [ˈjen.tɾəs] 'go-in.2SG.PRS.IND there'

Often hiatus between an unstressed vowel and a following stressed vowel follows from the *CLASH constraint. And while in stressed + unstressed sequences (§4.2.1, e.g. *destí incert* [dəs.ˈtin.ˈsɛrt]) clashes with one mora separating nuclei could be enforced due to the dominant *OCP/ˈV.V constraint, *CLASHμ prevented fusion or elision in *destí idoni* [dəs.ˈti .i.ˈðɔ.ni]. The same happens with unstressed + stressed vowel sequences. The asterisked forms in (22a) violate *CLASHμ. In (22b) the asterisked form violates *CLASH.

(22) (a) *territori inca* 'Inca territory' tə.ri.ˈtɔ.ɾi .ˈiŋ.kə *tə.ri.ˈtɔ.ˈɾiŋ.kə
foc i ambre 'fire and amber' ˈfɔ.ki .ˈam.bɾə *ˈfɔ.ˈkjam.bɾə
(b) *parli ara* 'speak.3SG.PRS.SUBJ now' ˈpar.li .ˈa.ɾə *ˈpar.ˈlja.ɾə

On the basis of the constraints considered up to this point, however, glide formation or elision ought to be available when the unstressed vowel stands in a proparoxytone word. Here *CLASH is not involved. Examples (23) from Bonet & Lloret (1998: 182) and Oliva (1980: 33) confirm that glide formation is ungrammatical here.

(23) *síntesi única* 'unique synthesis' *[ˈsin.tə.ˈzju.ni.kə] ≺ [ˈsin.tə.zi .ˈu.ni.kə]
ímpetu ari 'Aryan impetus' *[ˈim.pə.ˈtwa.ɾi] ≺ [ˈim.pə.tu .ˈa.ɾi]
pròtesi ampla 'broad prosthesis' *[ˈpɾɔ.tə.ˈzjam.plə] ≺ [ˈpɾɔ.tə.zi. ˈam.plə]
àrea única 'sole area' [ˈa.ɾe.ə .ˈu.ni.kə][4]
tètrica ombra 'dismal shade' [ˈtɛ.tɾi.kə. ˈom.bɾə]
càlida era 'hot threshing-floor' [ˈka.li.ðə. ˈe.ɾə]

It is conceivable that what is special about proparoxytone words is that they contain, in citation form, a well-formed disyllabic iambic foot to the right of the head foot: *síntesi* [ˈsin.tə.zi] (ˈσ) (σσ̲). (Here the head syllable of the non-head foot of the word is underlined.) The set of constraints established so far already has a Faithfulness-to-Prosodic-Heads (FPH) constraint ranked above anti-hiatus constraints, namely METRICAL CONSISTENCY(B,PPhr) (METCONSB-PHR): The head of a head foot in a base corresponds to the head of a syllable in a phonological phrase (11). The reason for specifying head foot in METCONSB-PHR was to distinguish preservation of initial [u] in *hindú honrat* (7a), which corresponds to a stressed vowel in the base, from the loss of initial [i] in

[4] This example is from Palmada (1994a: 127). In the present account there is no special phenomenon observed in examples like this, in which the final syllable of the first word lacks an onset.

destí incert (3b), where the [i] of *incert* is head of a bimoraic foot ([$i_{\mu}n_{\mu}$.]) which does not correspond to a stressed vowel (head foot) in the base. Preservation of *some* foot–head element seems to be involved in the pattern of (23). What is preserved in correspondence with the head of a non-head foot, though, is not precisely nuclearity (head of syllable) but rather moraicity. This can be seen in the example *menú imprès* ((1b) = (6b)) [mə'nujm.'pɾɛs], where the input syllable and unstressed foot head [i] retains its mora in coda position. In the cases of (23) retention of the mora of the final syllable of the proparoxytones preserves the vowel from becoming an onset (in the case of [i] or [u]) or from deletion (in the case of [ə]). But the example *jardí antic* [ʒəɾ.'ðin.'tik] 'old garden' (15) shows that it is not necessarily the *head* mora of a non-head (unstressed) foot that must be retained. In *jardí antic* the head mora of the non-head foot of *antic* [$ə_{\mu}n_{\mu}$.'tik] (σ)('σ) is lost in sandhi. However, in cases of resyllabification across a word boundary, the coda mora is lost when the coda becomes an onset, as, for example in *síntesis amples* 'broad syntheses' whose word inputs and phrase output are set out in (24). The input coda and its mora are in bold.

(24)

síntesis	*amples*	*síntesis amples*
'$sin.tə.zi_{\mu}\mathbf{s}_{\mu}$	'am.pləs	'$sin.tə.zi_{\mu}$.'**z**am.pləs

The examples seem to show, therefore, that what is preserved in correspondence involving an unstressed foot is *some* mora of the head syllable of the foot. This is expressed in the constraint in (25). The constraint need not mention non-head foot specifically, since the metrical consistency of head (stressed) feet is preserved in relevant respects by unviolated word-phrase faithfulness to prosodic heads (FPHPWd-PPhr).

(25) METRICAL CONSISTENCYPWd-PPhr (Foot mora) (METCONSW-P$\underline{\sigma}\mu$): Some mora of the head syllable of a foot in a prosodic word has a correspondent in a phonological phrase.

When a non-high unstressed vowel ([ə]) precedes a stressed one, as before, elision is blocked when a clash (one mora or none intervening) between phonological phrase heads would result, as in the examples of (26).

(26) *foca àrtica* ['fo.kə .'ar.ti.kə] 'arctic seal' (*foca*]$_{N}$ *àrtica*]$_{A}$)
rega arbres del jardí ['rɛ.ɣə .'a.βɾəz .ðəl .ʒəɾ.'ði] 'water.3SG.PRS.IND trees in the garden' (*rega*]$_{V}$ *arbres*]$_{N}$ *del jardí*]$_{PP}$)
assenyalar-ne una [ə.sə.ɲə.'lar.nə .'u.nə] 'point-out.INF one of-them'
assenyalar-ne altres [ə.sə.ɲə.'lar.nə .'al.trəs] 'point-out.INF others of-them'

The cases of [ə.'a] show why *'V_{Low}.ə (17) needs to specify the sequence 'stressed vowel + schwa'. While *'V_{Low}.ə ranks high—hence *mesurar alçària* [məzu'ɾal.'saɾjə] (12a) has a tolerated *CLASH violation—the more general constraint *OCP/V_{A}, which includes *ə.'a, ranks below *CLASH.

When the first word is in specifier position, or is a preposition (thus not a Phonological Phrase head and not protected by CLASH constraints), elision is normal (28a), reflecting the domination of ONSET over the correspondence constraint (27) MAXPWd-PPhr(ə).

(27) MAXPWd-PPhr(ə): [ə] in a prosodic word corresponds to [ə] in a phonological phrase.

(28) (a) *aquella època* [əˈkɛ.ˈʎɛ.pu.kə] 'that period'
la primera hora [lə.pɾi.ˈme.ˈɾɔ.ɾə] 'the first hour'
onze anys [ˈon.ˈzaɲs] 'eleven years'
entre altres [ˈen.ˈtɾal.tɾəs] 'among others'
(b) *l'escola era lluny* [ləs.ˈkɔ.ˈle.ɾə.ˈʎuɲ] 'the school was far' (subject, verb and adverb complement restructured as one phonological phrase)
ja no puja aigua del pou [ˈʒa. ˈno .ˈpu.ˈʒaj.ɣwə .ðəl .ˈpow] 'not raise. 3SG.PRS.IND water from the well any more' (verb *puja* and direct object *aigua* restructured into one PPhr)
quina és la millor? [ˈki.ˈnez .lə .mi.ˈʎo] 'which-one is the best?' (verb *és* incorporated within VP as PPhr: cf. *quina és?* [ˈki.nə .ˈes] 'which-one is it?')

In 'restructured' phonological phrases such as those of (28b) (see §9.6, and at (2) in §4.2.1), an incorporated element ceases to be a phonological phrase head, and then does not block elision. Palmada (1994a: 127) claims that elision of prevocalic [ə] is blocked when a vowel immediately precedes it, as in *via única* [ˈbi.ə .ˈu.ni.kə] 'single way', *tia índia* [ˈti.ə .ˈin.djə] 'Indian aunt', *insinua himnes* [in.si.ˈnu.ə .ˈim.nəs] 'insinuate.3SG.PRS.IND hymns'. But the relevant constraint here is *CLASH between prosodic phrase heads, just as in the examples of (26). Only her example *àrea única* [ˈa.ɾe.ə .ˈu.ni.kə] 'sole area' shows anything different. The relevant constraint blocking elision in this type of example is METRICAL CONSISTENCYPWd-PPhr (Foot mora); see (23), (25) above.

4.4 CONTACT BETWEEN UNSTRESSED VOWELS

Having dealt with sequences of vowels one of which is stressed (the first, §4.2, or the second, §4.3), I now consider contact between unstressed vowels. When unstressed vowels come into contact, the *CLASH constraint is not relevant. Reducing two unstressed syllables to one unstressed syllable can never lead to a *CLASH violation. However, reducing two unstressed syllables to one would often lead to a more harmonic outcome as evaluated by other constraints considered so far in this chapter. In fact, when each of the adjacent vowels is [ə], elision/fusion is the regular result, as in *perdre alè* [ˈpɛɾ.ðɾə.ˈlɛ] 'lose.INF breath',

escriptora aguda [əs.krip.ˈto.ɾə.ˈɣu.ðə] 'penetrating writer.F', with elision/fusion reflecting the effect of the ranking of the *OCP/V_A constraint (§2.3.9 (53)), or simply ONSET, above MAXPWd-PPhr(ə). But with pairs of identical high vowels—[i]+[i], [u]+[u]—patterns of hiatus or fusion/elision depend in part on the constraints already seen, such as members of the *LAPSE family, and in part on a constraint whose effect has not previously been evident. This new constraint is an output-output correspondence constraint against deleting palatal or labial vowels, identified here just as a MAX constraint (29) though a matching UNIFORMITY[5] constraint is also relevant in cases where absence of a vowel might be interpreted as fusion.

(29) MAXPWd-PPhr(I, U) (MAXW-P(I,U))

(30) (a) *medi inhòspit* 'inhospitable medium' [ˈmɛ.ði .i.ˈnɔs.pit] ≻ *[ˈmɛ.ði. ˈnɔs.pit]
porxo humit 'damp arcade' [ˈpɔr.ʃu .u.ˈmit] ≻ *[ˈpɔr.ʃu.ˈmit]
(b) *carro onzè* 'eleventh wagon', cf. *onze* 'eleven' [ˈka.run.ˈzɛ]
territori incaic 'Inca territory', cf. *inca* 'Inca' [tə.ri.ˈtɔ.ɾiŋ.ˈkajk]
soci irlandès 'Irish member' [ˈsɔ.sir.lən.ˈdɛs]
(c) *demani hipoteques* 'request.3SG.PRS.SUBJ mortgages' [də.ˈma.ni.pu. ˈtɛ.kəs]
piano holandès 'Dutch piano' [pi.ˈa.nu.lən.ˈdɛs]
(d) (= (3c)) *tabú oblidat* 'forgotten taboo' [tə.ˈβu.βli.ˈðat] ≻ *[tə.ˈβu. uβli.ˈðat]

The examples in (30a) show that elision/fusion is blocked in an open syllable,[6] while (30b) shows it is permitted in a closed syllable, even when corresponding to a stressed base. Examples in (30c) demonstrate the effect of the *LAPSE constraints. Example (30d) is introduced again to point up the contrast with other input [u.u] sequences in (30). Candidate evaluation for selected examples from (30) is illustrated in tableau (31).

[5] UNIFORMITY is the 'no fusion' constraint of correspondence theory (McCarthy & Prince 1995; 1999).

[6] The examples and the 'rule' are from Recasens (1993). Nevertheless, the transcribed texts in Veny and Pons (1998) contain examples like *quasi igual* [ˈka.zi.ˈɣɔl] 'almost the same', *oli i prou* [ˈɔ.li.ˈprɔw] 'oil and that's it', *un quilo o dos* [uŋ.ˈki.lu.ˈðos] 'one or two kilos' (with the /u/ variant of *o* 'or'). Not all speakers rank MAXW-P(I,U) higher than ONSET, evidently.

(31)

		*LAPSE3σ	*LAPSE3μ	METCONSB-Phr	*OCP/ˈV.V	MAXW-P(I,U)	ONSET
(a)	☞ ˈmɛ.ð̞i .i.ˈnɔs.pit						*
	ˈmɛ.ð̞i.ˈnɔs.pit					*!	
	☞ ˈpɔr.ʃu .u.ˈmit						*
	ˈpɔr.ʃu .ˈmit					*!	
(b)	ˈka.ru$_{\mu}$ .u$_{\mu}$n$_{\mu}$. ˈzɛ		*!				*
	☞ ˈka.run.ˈzɛ			*		*	
	tə.ri.ˈtɔɾi$_{\mu}$ .i$_{\mu}$ŋ$_{\mu}$.ˈkajk		*!				*
	☞ tə.ri.ˈtɔɾiŋ.ˈkajk			*		*	
	ˈsɔ.si$_{\mu}$ .i$_{\mu}$r$_{\mu}$.lə$_{\mu}$n$_{\mu}$.ˈdɛs		*!				*
	☞ ˈsɔ.sir.lən.ˈdɛs					*	
(c)	də.ˈma.ni$_{\mu}$ .i$_{\mu}$.pu$_{\mu}$.ˈtɛ.kəs		*!				*
	☞ də.ˈma.ni.pu.ˈtɛ.kəs					*	
	pi.ˈa.nu$_{\mu}$. u$_{\mu}$.lə$_{\mu}$n$_{\mu}$.ˈdɛs	*!	*				*
	☞ pi.ˈa.nu$_{\mu}$.lə$_{\mu}$n$_{\mu}$.ˈdɛs		*			*	
(d)	tə.ˈβu .u.βli.ˈð̞at				*!		*
	☞ tə.ˈβu .βli.ˈð̞at					*	

*LAPSE3σ » *LAPSE3μ » METCONSB-Phr, *OCP/ˈV.V » MAXW-P(I,U) » ONSET

When two unstressed high vowels of different quality are adjacent ([i]+[u], [u]+[i]), a falling diphthong is normal (32)—a rising diphthong is disfavoured by the *Cj, *Cw constraint mentioned, with respect to word-internal contexts, in §3.2.4.1 (and followed up in §3.2.4.8 and §3.2.5). There is little to be said here about constraint ranking, provided, of course, that *Cj, *Cw outranks *CODA.

(32) *porti olives* [ˈpɔr.tiw.ˈli.βəs] *[ˈpɔr.tju.ˈli.βəs] 'bring.3SG.PRS.SUBJ olives'
soci humil [ˈsɔ.siw.ˈmil] *[ˈsɔ.sju.ˈmil] 'humble member'
toro immens [ˈtɔ.ɾujm.ˈmens] *[ˈtɔ.ɾwim.ˈmens] 'huge bull'

When an unstressed high vowel ([i] or [u]) precedes an unstressed non-high vowel ([ə]), three possible outcomes need to be considered: hiatus ([i.ə], [u.ə]), a rising

diphthong ([jə], [wə]), and elision of the non-high (unmarked/default) vowel [ə]. Examples are given in (33). For each case, often two of the three alternatives occur in variation, while the third is ungrammatical. Where hiatus and a rising diphthong are both acceptable, hiatus generally belongs to a more formal or conservative style of pronunciation (reflecting relatively high ranking of *Cj,*Cw, as also word-internally; see §3.2.4.1, §3.2.4.8, §3.2.5). As elsewhere in this section, this optional hiatus is not specifically accounted for. (It will follow from higher-ranked UNIFORMσ or *Cj, *Cw ranked below ONSET.)

(33) (a) *compro anís* [ˈkom.pɾwə.ˈnis] (∼[ˈkom.pɾu .ə.ˈnis]) ≻ *[ˈkom.pɾu.ˈnis] 'buy.1SG.PRS.IND anisette'
tesi aguda [ˈtɛ.zjə.ˈɣu.ðə] (∼[ˈtɛ.zi .ə.ˈɣu.ðə]) ≻ *[ˈtɛ.zi.ˈɣu.ðə] 'penetrating thesis'
s'hi amaga [sjə.ˈma.ɣə] (∼[si. ə.ˈma.ɣə]) ≻ *[si.ˈma.ɣə] 'REFL-LOC hide.3SG.PRS.IND'
(b) *gerro estret* [ˈʒɛ.rus.ˈtɾɛt] (∼[ˈʒɛ.rwəs.ˈtɾɛt]) 'narrow jug'
pati esplèndid [ˈpa.tis.ˈplɛn.dit] (∼[ˈpa.tjəs.ˈplɛn.dit]) 'splendid courtyard'
pinto esglésies [ˈpin.tuz.ˈɣle.zjəs] (∼[ˈpin.twəz.ˈɣle.zjəs]) 'paint. 1SG.PRS.IND churches'
(c) *acullo artistes* [ə.ˈku.ʎwər.ˈtis.təs] (∼[ə.ˈku.ʎu .ər.ˈtis.təs]) ≻ *[ə.ˈku.ʎur.ˈtis.təs] 'welcome.1SG.PRS.IND artists'; cf. *art* 'art'
mesuri alçària [mə.ˈzu.ɾjəl.ˈsa.ɾjə] (∼[mə.ˈzu.ɾi .əl.ˈsa.ɾjə]) ≻ *[mə.ˈzu.ɾil.ˈsa.ɾjə] 'measure.3SG.PRS.SUBJ height'; cf. *alt* 'high'
(d) *dibuixo angelets* [di.ˈβu.ʃwənʲ.ʒə.ˈlɛts] (∼[di.ˈβu.ʃunʲ.ʒə.ˈlɛts]) ≻ *[di.ˈβu.ʃu .ənʲ.ʒə.ˈlɛts] 'draw.1SG.PRS.IND angel.DIM.PL'; cf. *àngel* 'angel'
porto americana [ˈpɔr.twə.mə.ɾi.ˈka.nə] (∼[ˈpɔr.tu.mə.ɾi.ˈka.nə]) 'wear.1SG.PRS.IND a jacket'

In (33a) it is seen that after an unstressed high vowel, initial [ə] in a pretonic open syllable may not be elided. Examples in (33b) show that elision is possible in an initial pretonic closed syllable, provided it does not correspond to the vowel of a stressed base, as seen in (33c). For the (33c) cases, the correspondence constraint METCONSB-Phr (11) comes into play again. Insofar as elision is preferred to a rising diphthong in (33b), as it appears to be in many varieties, this would reflect the general constraint against rising diphthongs, other things being equal. The other constraint that makes itself evident is a version of METCONSW-P$\underline{\sigma}$μ (25), relativized to a different prominent position, namely METCONSW-PInitσμ, which preserves some mora from an *initial* syllable against deletion or onset glide formation.[7] Thus (33a) shows METCONSW-PInitσμ

[7] METCONSW-PInitσμ is the nearest parallel in this account to Jiménez's Syllable Preservation Principle (1999: 97). This Syllable Preservation Principle, an output–output correspondence constraint, does not require that syllables be preserved absolutely: they may overlap as long as there remains some distinct independent material belonging to each.

» *Cj, *Cw, and ONSET » MAXW-Pμ; (33b) shows the effect of *Cj, *Cw » MAXPWd-PPhr(Vowel); (33c) shows METCONSB-Phr » *Cj, *Cw; and (33d) shows the dominance of *LAPSE constraints. The alternative versions in (33b) show MAXPWd-PPhr(Vowel) » *Cj, *Cw. (The acceptable variants in (33d) seem to show special disfavouring of rising diphthongs in a sequence of unstressed syllables, an issue which is not taken up further here.) These points are displayed in tableau (34).

(34)

		*LAPSE3μ	METCONSB-Phr	METCONSInitσμ	MAXW-P(I,U)	ONSET	*Cj, *Cw	MAXW-P(V)
(a)	☞ 'tɛ.zjə.'ɣu. ðə						*	
	'tɛ.zi.'ɣu.ðə			*!				*
	'tɛ.zi. ə'ɣu.ðə					*!		
(b)	☞ 'pɔr.ʃu .u.'mit					*		
	'pɔr.ʃu.'mit				*!			*
(c)	☞ 'ʒɛ.rus.'trɛt							*
	'ʒɛ.rwəs.'trɛt						*!	
	'ʒɛ.ru .əs.trɛt					*!		
(d)	☞ ə.'ku.ʎwər.'tis.təs						*	
	ə.'ku.ʎur.'tis.təs		*!					*
	ə.'ku.ʎu .ər.'tis.təs	*!				*		

*LAPSE3μ » METCONSB-Phr » METCONSInitσμ,
MAXW-P(I,U) » ONSET » *Cj, *Cw » MAXW-P(V)

When a non-high unstressed vowel ([ə]) precedes a high one, a falling diphthong ([əj], [əw]) or elision of [ə] are both possible (35). Bonet & Lloret (1998: 185) remark that elision is commoner, for example, in the Girona region than in the Barcelona region.[8]

[8] A regional difference within eastern Catalan is not clearly evident in the texts of Veny & Pons (1998). In those texts, elision dominates glide formation approximately in the proportion 4:1. In some 73% of cases the high vowel involved is in the indefinite article *un, una* (as in *agafava un drap* [əɣəˈfa.βunˈdrap] 'took a cloth'), but this does not seem especially significant, as in the remaining examples (e.g. *roba interior* [ˈrɔ.βin. tə.ˈɾjor] 'underwear') the ratio of elision to glide formation is 7:2. In establishing this count I have not taken *em*, *et*, *es*, *el*, *ens*, *els*, orthographic forms of pronominal clitics, to be underlyingly vowel-initial.

(35)	*porta ulleres*	ˈpɔr.təw.ˈʎe.ɾəs ∼ ˈpɔt.tu.ˈʎe.ɾəs
	‘wear.3SG.PRS.IND glasses’	(ˈσ) (σ) (ˈσ) σ (ˈσ)(σ ˈσ) σ
	cosa inútil ‘useless thing’	ˈkɔ.zəj.ˈnu.til ∼ ˈkɔ.zi.ˈnu.til
		(ˈσ)(σ)(ˈσ) σ (ˈσ)(σ ˈσ)σ
	aquesta imatge ‘this image’	ə.ˈkɛs.təj.ˈmad͡dʒə ∼ ə. ˈkɛs.ti.ˈmad͡dʒə
		(σ ˈσ)(σ) (ˈσ) σ (σ ˈσ) (σ ˈσ) σ
	queda oberta	ˈke.ðəw.ˈβɛr.tə ∼ ˈke.ðu.ˈβɛr.tə
	‘stay.3SG.PRS.IND open’	(ˈσ)(σ) (ˈσ) σ (ˈσ)(σ ˈσ) σ
	demana-ho ‘request.2SG.IMP it’	də.ˈma.nəw ∼ də.ˈma.nu
		(σ ˈσ) (σ) (σ ˈσ) σ
	creure-hi ‘believe.INF in-it’	ˈkɾɛw.ɾəj ∼ ˈkɾɛw.ɾi
		(ˈσ)(σ) (ˈσ) σ
	la universitat ‘the university’	ləw.ni.βər.si.ˈtat ∼ lu.ni.βər.si.ˈtat
		(σ) (σ σ) (σ ˈσ) = (σ)(σ σ)(σ ˈσ)

A question that may arise in this context is why elision exists at all as an option in (35), since glide formation is sufficient to avoid ONSET violation, as it does word-internally. Elision additionally violates some version of MAX. It appears that the elision option is favoured by rhythmic considerations such as might reflect a *STRUCTURE constraint as indicated in the footings in the right-hand column of (35); elision/fusion results in more light syllables and fewer feet.

4.5 OBSERVATIONS ON MULTI-VOCOID SEQUENCES

Though Recasens (1993: 139–44) devotes several pages to sandhi involving sequences of more than two vocoids, there is, perhaps unsurprisingly, relatively little that does not fall out directly from the constraints already discussed. It may be worthwhile to draw attention to a few matters of special interest. In §3.2.5.2 the issue of the interpretation and realization of sequences of identical high vocoids was raised. It was remarked there that this question prompts a distinction between a production grammar and a perception grammar, and evidence here from vowel sandhi supports this perspective.

(36) (a) *mateu una gallina* [mə.ˈtɛ. u.nə.ɣə.ˈi.nə] ‘kill.2PL.PRS.IND a chicken’ (/maˈtɛw/+/una/) *[mə.ˈtɛw.nə]

deu o dotze [ˈðɛ.u.ˈðod͡.dz͡ə] ‘ten or twelve’ (/ˈdɛw/+/u/) *[ˈðɛw.ˈðod͡.dz͡ə]

allò se’n diu una perdigussa [ə.ˈʎɔ.sən.ˈdi .u.nə.pər.ðiˈɣu.sə] ‘that is known as a *perdigussa* (partridge trap)’ (/ˈdiw/ + /una/) *[ˈdiw.nə][9]

tu ho ofereixes [tu.u.fə.ˈɾɛ.ʃəs] ‘you.SG offer.2SG.PRS.IND it’ (/ˈtu/ + /u/ + /ofeɾ+ˈɛʃ+s/; the morpheme /u/ is not underlyingly specified for syllabicity) *[tuw.fə.ˈɾɛ.ʃəs], *[tu.fə.ˈɾɛ.ʃəs]

[9] The above three examples are from the eastern continental texts in Veny & Pons (1998).

jo hi imagino [jɔ.i.mə'ʒi.nu] 'I imagine.1SG.PRS.IND there' (/'jɔ/ + /i/ + /ima'ʒin+o/; the morpheme /i/ 'there' is not underlyingly specified for syllabicity) *[jɔj.mə'ʒi.nu]

(b) *el carro ho interceptava* [əl.'ka.ru.in.tər.səp.'ta.βə] 'the cart intercept. 3SG.PST.IMPF it' (/'kaRRo/ + /u/ + /inter-/) *['ka.rujn.tər.səp.'ta.βə]

presideixi i actui [prə.zi.'ðɛ.ʃi .ək.'tu.i] 'preside.3SG.PRS.SUBJ and act.3SG.PRS.SUBJ' (/pRezid+'ɛʃ+i/ + /i/ + /ak'tu+i/; the morpheme /i/ 'and' is not underlyingly specified for syllabicity) *[prə.zi.'ðɛ.ʃjək.'tu.i], cf. *presideix i actua* [prə.zi.'ðɛ.ʃjək.'tu.ə] 'preside3SG.PRS.IND and act.3SG.PRS.IND'

In the examples in (36a) the input involves in each case a sequence involving a vowel followed by two high vocoids matching in place of articulation. From the perspective of the production grammar, the first of the two high vocoids is syllabified either as a glide in the input prosodic word or as a glide by default (between two vowels, given ONSET and *CODA). However, from the perception viewpoint [wu], [ji] cannot be adequately distinguished from [u], [i] respectively, which is how the output is transcribed. But if [wu], [ji] were completely indistinguishable from [u], [i] in both production and perception, we should expect onsetless [.u] and [.i] themselves to be subject to glide formation. The examples show, on the contrary, that [u], [i], representing /w+u/, /j+i/, are not subject to glide formation: they are not treated like onsetless syllables, despite the fact that they 'sound' like onsetless syllables. The issue is essentially the same in (36b); intended prevocalic /u+w/, /i+j/, cannot be adequately distinguished, as far as perception is concerned, from [u] and [i]. In these cases (36b) Recasens offers transcriptions with and without glides as alternatives ([uwV] ∼ [uV], [ijV] ∼ [iV]), though he retains the glide consistently when it is word-initial, as in *faci ioga* ['fa.si.'jɔ.ɣə] 'do.3SG.PRS.SUBJ yoga'. The latter transcription practice no doubt reflects the especial psychological salience of a word-initial onset, even when, perceptually, it bears no features distinguishing it from the preceding nucleus. The functional morphemes *ho* [u] ∼ [w] 'it', *o* [u] ∼ [w] (but not the stressed variant ['ɔ]) 'or', *hi* [i] ∼ [j] (locative pronominal clitic), and *i* [i] ∼ [j] 'and' are here understood to have no inherent or underlying specification for syllabicity (see §§3.2.5.1 and 3.2.4 for arguments that some high vocoids must be so specified in Catalan.)

The realization of word sequences confirms the ranking ONSET » MAXW-P(ə) (§4.3 (27)) or the activity of *'V_{Low}.ə (§4.2.2 (13)) or *OCP/V_A (§2.3.9 (53)).

(37) *va passar a obtenir bons resultats* ['ba.pə.'sawp.tə.'ni . . .] 'AUX.3SG. PST.PRF go-on to get good results'

una casa a Andalusia [u.nə. 'ka.zən.də.lu.'zi.ə] 'a house in Andalusia'

As the examples in (37) show, there is no barrier to the preposition *a* 'to' disappearing altogether, when it stands adjacent to ['a] or another [ə], reflecting the ranking ONSET » MAXPPhr-PrWd(ə). As the examples at (28) showed, elision

of final [ə] is the norm before a stressed vowel, although, as was observed there, elision can be blocked by *CLASH. Elision of [ə] before an unstressed vowel is also common (examples (33) and (35)). Elision of final [ə] may also, in a narrow range of contexts, be blocked by METRICALCONSISTENCYPWd-PPhr (Foot mora) (§4.3, (23), (25)). This is where Palmada's (1994a: 127) example *àrea única* [ˈa.ɾe.ə .ˈu.ni.kə] 'sole area' fits in. Metrical consistency constraints also contribute to preserving one [ə] in a vowel sequence in the example *conrea aïlladament* [kun.ˈre.ə.i.ˈʎa.ðə.ˈmen] 'cultivate.3SG.PRS.IND in-isolation'. A range of candidates for this phrase is evaluated in tableau (38) by a number of the relevant constraints considered in this section.

(38)

	conrea aïlladament cf. [ə.ˈi.ʎə] 'isolate.3SG.PRS.IND'	*LAPSE3μ	METCONSB-Phr	METCONSInitσμ	*OCP/VA	ONSET
(a)	kun.ˈre.ə.ə.i.ˈʎa.ðəˈmen	*!			*	***!
(b)	☞ kun.ˈre.ə.i.ˈʎa.ðə.ˈmen					**
(c)	kun.ˈre.i.ˈʎa.ðəˈmen			*!		*
(d)	kun.ˈre.əj.ˈʎa.ðəˈmen		*!			*
(e)	kun.ˈrej.ˈʎa.ðə.ˈmen		*!	*		

In (38) the most harmonic candidate is (38b), with two ONSET violations. In fact, candidate (38c) is also acceptable, for Recasens. This may reflect another stylistically variable rhythmic *LAPSE constraint, specifically *LAPSE2σ, that would rank above METCONSINITσμ. Recasens, following Palmada (1994a), considers the pronunciation involving words like *àrea* only in phrases in which they are followed by a stressed initial vowel. Further investigation is needed for *àrea* (or words of similar structure) followed by unstressed vowel-initial words of various length, such as those in (39). Native speakers I have consulted are rather hesitant about their preferences, though elision of the final [ə] of *àrea* seems disfavoured, as the position of METCONSσμ in this grammar would predict.

(39) *àrea impossible* 'impossible area' (initial heavy, not corresponding to a stressed base, followed by an unstressed syllable, i.e. with an input four-syllable, five-mora lapse)
àrea inhòspita 'inhospitable area' (initial light, not corresponding to a stressed base, pretonic, i.e. with an input three-syllable, three-mora lapse)
àrea infecta 'tainted area' (initial heavy, not corresponding to a stressed base, pretonic, i.e. with an input three-syllable, four-mora lapse)
àrea idòlatra 'idolatrous area', cf. *ídol* 'idol' (initial light, corresponding to a stressed base, pretonic, i.e. with an input three-syllable, three-mora lapse)

àrea incaica 'Inca area', cf. *inca* 'Inca' (initial heavy, corresponding to a stressed base, pretonic, with an input three-syllable, four-mora lapse)

Lastly, consider a few examples of unstressed high–non-high–high sequences as in (40).

(40) *porto auriculars* [ˈpɔr.tu.ɾi.ku.ˈlars] ~ [ˈpɔr.twəw.ɾi.ku.ˈlars] 'wear.1SG.PRS.IND headphones'
a la sàrria hi posaven [ə.lə.ˈsa.ri.pu.ˈza.βən] ~ [ə.lə.ˈsa.rjəj.pu.ˈza.βən] 'in the saddle-bag LOC put.3PL.PST.IMPF'
desitjo airejar-me ~ [də.ˈzi͡d.d͡ʒuj.ɾə.ˈʒar.mə] ~ [də.ˈzi͡d.d͡ʒwəj.ɾə.ˈʒar.mə] 'wish.1SG.PRS.IND to take the air'; cf. *aire* 'air'

The shorter alternatives in (40) are, as before, the ones favoured when the allegro style *LAPSE3μ rhythm constraint is active. Tableau (41) shows how this works, indicating in parentheses the winner when *LAPSE3μ is not active. That (41a.ii) is the winner when *LAPSE3μ is ignored shows, by the way, that UNIFORMITY (No fusion) outranks ONSET.

(41)

		*LAPSE3μ	METCONSB-Phr	METCONSInitσμ	MAXW-P(I, U)	UNIFORMITY	ONSET	*Cj, *Cw
(a)	*porto*$_1$ *a*$_2$*u*$_3$*ricular*s							
	i. ˈpɔr.tu.əw.ɾi.ku.ˈlars	μμ!					*(!)	
	ii. (☞) ˈpɔr.twəw.ɾi.ku.ˈlars	μμ!						*
	iii. ☞ ˈpɔr.tu$_{13}$.ɾi.ku.ˈlars	μ				*(!)		
(b)	*sàrr*$_1$*ia*$_2$ *hi*$_3$ *posaven*							
	i. (☞) sa.rjəj.pu.ˈza.βən	μ!						*
	ii. ☞ ˈsa.ri$_{13}$.pu.ˈza.β nə					*(!)		
(c)	*desitjo airejar-me*							
	i. də.ˈzi͡d.d͡ʒu.əj.ɾə.ˈʒar.mə	μμ!					*	
	ii. ☞ də.ˈzi͡d.d͡ʒwəj.ɾə.ˈʒar.mə	μ						*
	iii. də.ˈzi͡d.d͡ʒu j.ɾə.ˈʒar.mə	μ	*!					

*LAPSE3μ » METCONSB-Phr, METCONSInitσμ, MAXW-P(I, U), UNIFORMITY » ONSET » *Cj, *Cw

In (41c), with or without *LAPSE3μ, the winner with the constraint ranking we are considering is (41c.ii), as (41c.iii) violates METCONSB-Phr. The acceptability of candidate (41c.iii) results from ignoring the morphological connection with *aire* [ˈajɾə] 'air' that METCONSB-Phr acknowledges.

4.6 SUMMARY OF EVIDENCE FOR CONSTRAINT RANKINGS FOR VOWEL SANDHI

Constraint ranking	Winner and loser	
ONSET » UNIFORMσ	mə.ˈnujm.ˈpɾɛs ≻ mə.ˈnu. im.ˈpɾɛs	*menú imprès* ‘printed menu’ (1b) = (6b)
*Cj, *Cw » MAXW-P(ə)	ˈʒɛ.rus.ˈtɾɛt ≻ ˈʒɛ.rwəs.ˈtɾɛt	*gerro estret* ‘narrow jug’ (33b)
ONSET » *Cj, *Cw	ˈtɛz.jə.ˈɣu.ðə ≻ ˈtɛ.zi .ə.ˈɣu.ðə	*tesi aguda* ‘penetrating thesis’ (33a)
UNIFORMITY » ONSET	ˈpɔr.ʃu .u.ˈmit ≻ ˈpɔr.ʃu$_{12}$.ˈmit	*porxo humit* ‘damp arcade’ (30a)
MAXW-P(I,U) » ONSET	ˈpɔr.ʃu .u.ˈmit ≻ ˈpɔr.ʃu$_{1}$.ˈmit	*porxo humit* ‘damp arcade’ (30a)
*OCP/V$_A$ » UNIFORMITY	ˈpɛɾ.ðɾə.ˈlɛ ≻ ˈpɛɾ.ðɾə .ə.ˈlɛ	*perdre alè* ‘lose.INF breath’ §4.4
METCONSInitσμ » *Cj, *Cw	ˈtɛz.jə.ˈɣu.ðə ≻ ˈtɛ.zi.ˈɣu.ðə	*tesi aguda* ‘penetrating thesis’ (33a)
*CLASH » ONSET	mə.ˈnu .i.ˈðɔ.ni ≻ mə.ˈnuj.ˈðɔ. ni	*menú idoni* ‘suitable menu’ (1a) = (6a)
*CLASH » METCONSInitσμ	ˈpar.li .ˈa.ɾə ≻ ˈpar.ˈlja.ɾə	*parli ara* ‘speak.3SG.PRS.SUBJ now’ (22b)
*ˈV.ə » METCONSInitσμ	pəs.ˈto.lə.ˈmaɲ ≻ pəs.ˈto .ə.lə.ˈmaɲ	*pastor alemany* ‘German shepherd’ (17e)
*ˈV.ə » *CLASH	ʒər.ˈðin.ˈtik ≻ ʒər.ˈði .ən.ˈtik	*jardí antic* ‘old garden’ (15)
*OCP/ˈV.V » METCONSInitσμ	tə.ˈβu.βli.ˈðat ≻ tə.ˈβu .u.βli.ˈðat	*tabú oblidat* ‘forgotten taboo’ (3c)
*OPC/ˈV.V » *CLASH	dəs.ˈtin.ˈsɛrt ≻ dəs.ˈti .in.ˈsɛrt	*destí incert* ‘uncertain fate’ (3b)
METCONSB-Phr » *Cj, *Cw	ə.ˈku.ʎwər.ˈtis.təs ≻ ə.ˈku.ʎur.ˈtis.təs	*acullo artistes* ‘welcome.1SG.PRS.IND artists’ (33c)
METCONSB-Phr » *ˈV.ə	sən.ˈti .ə.βi.ˈðɛ.zə ≻ sən.ˈti.βi.ˈðɛ.zə	*sentir avidesa* ‘feel.INF greed’ (17b)
METCONSB-Phr » *OCP/ˈV.V	in.ˈdu .u$_{\mu}$n$_{\mu}$.ˈrat ≻ in.ˈdun.ˈrat	*hindú honrat* ‘honest Hindu’ (7a), (10)
METCONSσμ » ONSET	ˈa.ɾe.ə .ˈu.ni.kə ≻ ˈa.ɾe.ˈu.ni. kə	*àrea única* ‘sole area’ (23)
*LAPSE3μ » METCONSB-Phr	in.ˈdun.rə$_{\mu}$.ˈði.sim ≻ in.ˈdu .u$_{\mu}$n$_{\mu}$.rə$_{\mu}$.ˈði.sim	*hindú honradíssim* ‘most honest Hindu’ (7c), (10)
*LAPSE3μ » UNIFORMITY	ˈka.run.ˈzɛ ≻ ˈka.ru .un.ˈzɛ	*carro onzè* ‘eleventh wagon’ (30b)
*LAPSE3μ » *Cj, *Cw	ə.ˈku.ʎwər.ˈtis.təs ≻ ə.ˈku.ʎur.ˈtis.təs	*acullo artistes* ‘welcome. 1SG.PRS.IND artists’ (33c)
*CLASHμ » *OCP/ˈV.V	dəs.ˈti .i.ˈðɔ.ni ≻ dəs.ˈti.ˈðɔ.ni	*destí idoni* ‘suitable fate’ (3a)
*ˈV$_{\text{Low}}$.ə » *CLASHμ	ə.ˈʃɔ.ˈni.mə ≻ ə.ˈʃɔ .ə.ˈni.mə	*això anima* ‘that encourages’ (12b)
*INOV » *OCP/ˈV.V	ˈdu.ul.ˈtɾad͡dʒə ≻ ˈdul.ˈtɾad͡.dʒə	*dur ultratge* ‘serious insult’ (18a)
*LAPSE3σ » *INOV	ˈbiks.tɾəwɾ.ði.ˈna.ɾi ≻ ˈbi .əks.tɾəwɾ.ði.ˈna.ɾi	*vi extraordinari* ‘extraordinary wine’ (18d)

5

CODA VOICING NEUTRALIZATION AND ASSIMILATION

5.1 DISTRIBUTION OF VOICED AND VOICELESS OBSTRUENTS

In the phonological literature Catalan is quite well known as an example of a language which neutralizes voice contrasts among obstruents in coda position. In this section I review the facts of contrast and neutralization among obstruents in Catalan (Recasens 1993: 170–72; Bonet & Lloret 1998: 103–5, 118–26; Palmada 2002). In (1) and (2) are displayed examples of voicing contrasts. In (1) we see obstruents contrasting in voicing in prevocalic onset position.

(1)

		_V	
Manner	Place	Voiceless	Voiced
Non-strident	Labial	*pota* [ˈpotə] ‘paw’	*bota* [ˈbotə] ∼ [ˈβotə] ‘barrel’
	Dental	*tona* [ˈtonə] ‘ton’	*dóna* [ˈdonə] ∼ [ˈðonə] ‘woman’
	Velar	*cas* [ˈkas] ‘case’	*gas* [ˈɡas] ∼ [ˈɣas] ‘gas’
Strident fricative	Labio-dental	*fi* [ˈfi] ‘end’	*vi* [ˈvi] ‘wine’
	Alveolar	*sona* [ˈsɔnə] ‘sounds’	*zona* [ˈzonə] ‘zone’
	Alveolo-palatal	*xerra* [ˈʃɛrə] ‘chatters’	*gerra* [ˈʒɛrə] ‘jar’
Strident affricate	Alveolar	*potser* [puˈt͡se] ‘maybe’	*dotze* [ˈdod͡dzə] ‘twelve’
	Alveolo-palatal	*metxa* [ˈmet͡tʃə] ‘wick’	*metge* [ˈmed͡dʒə] ‘doctor’

Those of the contrasts that are illustrated in (1) in word-initial onset position are equally to be found in word-medial onset position. The /f/ – /v/ contrast is absent in most varieties, in which /v/ has merged with /b/ (§2.1.3.1). See §10.1 on the alternation between voiced stops and fricatives—essentially the fricatives occur in onsets between continuants ([+cont]).

(2)

		_liquid, _glide	
Manner	Place	Voiceless	Voiced
Non-strident	Labial	*prou* [ˈpɾɔw] ‘enough’	*brou* [ˈbɾɔw] ∼ [ˈβɾɔw] ‘broth’
		plau [ˈplaw] ‘pleases’	*blau* [ˈblaw] ∼ [ˈβlaw] ‘blue’
		sípia [ˈsipjə] ‘cuttlefish’	*ràbia* [ˈraβjə] ‘rage’
	Dental	*truita* [ˈtɾujtə] ‘trout’	*druida* [ˈdɾujðə] ∼ [ˈðɾujðə] ‘druid’
		llàntia [ˈʎantjə] ‘lamp’	*òndia!* [ˈɔndjə] ‘gosh!’
	Velar	*creu* [ˈkɾɛw] ‘cross’	*greu* [ˈɡɾew] ∼ [ˈɣɾew] ‘serious’
		classe [ˈklasə] ‘class’	*glaça* [ˈglasə] ∼ [ˈɣlasə] ‘freeze’
		qualla [ˈkwaʎə] ‘congeals’	*guatlla* [ˈɡwaʎʎə] ∼ [ˈɣwaʎʎə] ‘quail’
Strident fricative	Labio-dental	*fruita* [ˈfrujtə] ‘fruit’	*
		flor [ˈflɔ] ‘flower’	
		pífia [ˈpifjə] ‘blunder’	*núvia* [ˈnuvjə] ‘bride’

All of the preconsonantal obstruent voicing contrasts illustrated in (2) in initial onsets are also found in word-medial onsets. Notice, however that there is no contrast between /tl/ and /dl/. A medial sequence of coronal stop plus alveolar lateral exists in Catalan, but the stop must be voiced and the cluster is syllabified as coda + onset (e.g. *atleta* [əd.ˈlɛtə] ∼ [əl.ˈlɛtə] ‘athlete’). In the voiceless column parallel to the asterisked cell, onset /fr/ and /fl/ are found, and these also exist medially (e.g. *xamfrà* [ʃəmˈfra] ‘chamfer’, *bafle* [ˈbaflə] ‘speaker cabinet’); but the voiced fricative clusters are absent even from varieties which have /v/. This gap is usually interpreted as fortuitous—there is no source in the ancestor language, and a parallel gap exists in those languages from which Catalan lexical borrowings come.

In all other environments than in initial or medial onsets there is neutralization of voice contrasts in obstruents. This neutralization has major morphophonemic consequences when obstruents stand in word-final position (alternating with contrasting obstruents in prevocalic onset position). In (3) are examples of word-final neutralization of voicing in obstruents. The forms in the right-hand column, in which the stem is followed by a vowel-initial suffix, demonstrate the underlying voice value.

(3)		Neutralization / __ #	Contrast / __V
	/b/	*sap* [ˈsap] ‘knows’	*sabem* [səˈβɛm] ‘know.1PL.PRS.IND’
	/p/	*tip* [ˈtip] ‘fed up.M’	*tipa* [ˈtipə] ‘fed up.F’
	/d/	*verd* [ˈbɛɾt] ‘green.M’	*verda* [ˈbɛɾðə] ‘green.F’
	/t/	*fort* [ˈfɔrt] ‘strong.M’	*forta* [ˈfɔrtə] ‘strong.F’
	/g/	*fang* [ˈfaŋ(k)] ‘mud’	*fangós* [fəŋˈgos] ‘muddy’
	/k/	*ronc* [ˈroŋ(k)] ‘hoarse.M’	*ronca* [ˈroŋkə] ‘hoarse.F’
	/v/	*serf* [ˈserf] ‘serf.M’[1]	*serva* [ˈserβə] ∼ [ˈseɾvə] ‘serf.F’
	/f/	*buf* [ˈbuf] ‘puff’	*bufar* [buˈfa] ‘to puff’
	/z/	*cas* [ˈkas] ‘case’	*casos* [ˈkazus] ‘cases’
	/s/	*braç* [ˈbɾas] ‘arm’	*braços* [ˈbɾasus] ‘arms’
	/ʒ/	*roig* [ˈrot͡ʃ] ‘red.M’	*roja* [ˈroʒə] ‘red.F’
	/ʃ/	*calaix* [kəˈlaʃ] ‘drawer’	*calaixos* [kəˈlaʃus] ‘drawers’
	/dʒ/	*mig* [ˈmit͡ʃ] ‘half.M’	*mitja* [ˈmid͡ʒə] ‘half.F’
	/tʃ/	*despatx* [dəsˈpat͡ʃ] ‘office’	*despatxos* [dəsˈpat͡ʃus] ‘offices’

Examples of voice neutralization of obstruents in internal codas are given in (4); (4a) shows obstruents before non-strident obstruents, (4b) shows obstruents before strident obstruents (affricates as in (1) could well be included here—there are no [-tz-], [-ds-], [-tʒ-], or [-dʃ-] clusters); (4c) shows obstruents before nasals; and (4d) shows obstruents before onset liquids.

(4)	Neutralization / __ voiceless stop	Neutralization / __ voiced stop
(a)	*dubte* [pt] ‘doubt’	*súbdit* [bd] ‘subject’
		futbol [db] ‘football’
		addicció [dd] ‘addiction’
	acte [kt] ‘act’	*maragda* [gd] ‘emerald’
	diftong [ft] ‘diphthong’	
		afgà [vɣ] ‘Afghan’
	crespa [sp] ‘curly.F’	*bisbe* [zβ] ‘bishop’
	tastar [st] ‘to taste’	*esdrúixol* [zð] ‘proparoxytone’
	fosca [sk] ‘dark.F’	*pelasga* [zɣ] ‘Pelasgian’
	Neutralization / __ voiceless fricative	Neutralization / __ voiced fricative
(b)	*capsa* [ps] ‘box’	
		objecte [bʒ] ‘object’
		adverbi [dv] ∼ [db] ‘adverb’
	adscriure [tsk] ‘assign’	
	sexe [ks] ‘sex’	*èczema* [gz] ‘eczema’
	esfera [sf] ‘sphere’	*esvera* [zv] ∼ [zβ] ‘terrifies’

[1] This is the only example in the lexicon in which stem-final /v/ alternates with /f/, and it is also the only word in which stem-final /v/ is not preceded by a vowel. Post-vocalic /v/ alternates with /w/ in word-final position (§10.3).

	Neutralization / __ [n]	Neutralization / __ [m]
(c)	*hipnosi* [bn] 'hypnosis'	*submarí* [bm] 'submarine'
	ètnic [dn] 'ethnic'	*ritme* [dm] 'rhythm'
	signe [gn] 'sign'	*dracma* [gm] 'drachma'
	ovni [vn] 'UFO'	
	esnob [zn] 'snob'	*maresma* [zm] 'saltmarsh'

	Neutralization / __ [l]	Neutralization / __ [ʎ]	Neutralization / __ [r]
(d)	*legislar* [zl] 'to legislate'	*deslletar* [zʎ] 'to wean'	*Israel* [zr] ∼ [r] 'Israel'
	qui-sap-lo [bl] 'a lot'		*subratllar* [br] 'to underline'
	atleta [dl] 'athlete'		

Observe that in all of (4a–d) coda obstruents agree in voicing with a following onset consonant. (In (4c) coda stops are also variably subject to nasal assimilation, and coronal stops also to place assimilation; thus *ritme* may be pronounced [ˈrimmə]: see §6.3.) Exactly the same pattern of voicing assimilation as in (4) is also seen in sandhi consonant groups, in contiguous words, whether in phrases or in compounds (5). Observe that glides are included among the voiced consonants triggering voicing assimilation of final stops.

(5)	Neutralization / __ voiceless	Neutralization / __ voiced
	sap cantar [pk] 'knows how to sing'	*sap riure* [br] 'knows how to laugh'
	m'ha fet pregar [tp] ∼ [pp] 'has made me ask'	*m'ha fet llegir* [dʎ] ∼ [ʎʎ] 'has made me read'
	llarg camí [kk] 'long road'	*llarg de cames* [gd] 'long-legged'
		poc whisky [gw] 'not much whisky'
	buf calent [fk] 'warm puff'	*buf d'aire* [vð] 'puff of air'
	cas terrible [st] 'terrible case'	*cas notable* [zn] 'remarkable case'
	mateix fil [ʃf] 'same thread'	*mateix líquid* [ʒl] 'same liquid'
	mig cru [t͡ʃk] 'half-raw'	*migdia* [d͡ʒð] 'midday'

Lastly, when word-final obstruents are followed by vowel-initial words, non-sibilant plosives and fricatives are voiceless, but sibilants (fricatives and affricates) are voiced, independently of any underlying voice contrast in the word-final obstruent, as in (6). The elements of a word-final obstruent cluster agree in voicing (6b).

(6)	Voiceless non-sibilant obstruent / __ #V	Voiced sibilant / __ #V
(a)	*sap ajudar* [p] 'knows how to help'	*cas extrem* [z] 'extreme case'
	pot ajudar [t] 'may.3SG help'	*mateix element* [ʒ] 'same element'
	llarg any [k] 'long year'	*mig any* [d͡ʒ] 'half a year'
	xef únic [f] 'sole chef'[2]	
(b)		*cops amagats* [bz] 'hidden blows'
		pots ajudar [d͡z] 'may.2SG help'
		llargs anys [gz] 'long years'

Those varieties (*apitjat*) which lack voiced sibilant phonemes (see § 2.1.3.6) have voiceless word-final sibilants in the prevocalic contexts of (6); consequently, in those dialects the whole cluster in the (6b) examples is voiceless. However, these varieties do show assimilation of voicing in sibilants to following obstruents, as in (4) and (5). Varieties of southern Valencian (*alacantí*), and in some other areas such as southern Aragon, do not have voiceless plosives in word-final position before a vowel (6a), but rather show voiced non-strident fricatives in this context (Jiménez 1999: 176; Wheeler 1986): [ˈsaβ ad͡ʒuˈðaɾ], [ˈpɔð ad͡ʒuˈðaɾ], [ˈʎaɾɣ ˈaɲ].

5.2 'INCOMPLETE' NEUTRALIZATION

The word-final voicing neutralization presented in §5.1 might more accurately be described as *perceptual* neutralization. That is, native speakers do not perceive a difference between, for example, *cap* 'no' and *cap* /kab/ 'fit.3SG.PRS.IND', or between *ric* /ˈrik/ 'rich' and *ric* /ˈri+g/ 'laugh.1SG.PRS.IND'. There is modest evidence that some speakers, on some occasions, produce detectable differences between 'underlying' voiced and 'underlying' voiceless non-strident obstruents in word-final position, in experimental reading-style conditions. This evidence was first reported by Dinnsen & Charles-Luce (1984), and sparked a drawn-out controversy concerning both the experimental data and method and its interpretation (Mascaró 1987a; Charles-Luce & Dinnsen 1987; Charles-Luce 1993; Manaster-Ramer 1996a; Port 1996; Manaster-Ramer 1996b). The reanalysis of the first experiment involving 'perceptual homophones' such as those mentioned above (Charles-Luce & Dinnsen 1987) showed a mean advantage of underlying voice of 4 per cent in the duration of voicing into closure. The second experiment (Charles-Luce 1993) found an advantage in the length of the preceding vowel of 17 per cent for underlying voice in semantically neutral contexts, but of only 6 per cent in contexts semantically favouring the appropriate sense of the 'homophone'. Notice that although both studies detected differences associated

[2] Recasens (1993: 171) notes variable voicing of prevocalic word-final /f/ in positions distant from phrase accent, e.g. *biògraf alemany* [f] ~ [v] 'German biographer'.

with underlying voicing, the cues that reached statistical significance were different in the two experiments.

In a paper primarily addressing a different phonological issue, Dols & Wheeler (1996: 55–6) draw attention to the fact that, in Majorcan Catalan, obstruent voice contrasts may be maintained in certain unusual codas, namely, complex clusters involving a stop preceded by a nasal and followed by a liquid—clusters occurring only in first person singular present indicative verb forms that are characteristic of the Majorcan variety, as in (7a).

(7) Majorcan 1SG present indicative forms

	Voiceless	Voiced
(a)	*empr* [pr̥] 'I use'	*sembr* [bɾ] 'I sow'
	compr [pr̥] 'I buy'	*escombr* [bɾ] 'I sweep'
	concentr [tr̥] 'I concentrate'	*cilindr* [dɾ] 'I roll'
	entr [tr̥] 'I enter'	*emmangr* [gɾ] 'I paint with red ochre'
	umpl [pl̥] 'I fill'	*sembl* [bl] 'I seem'
	vincl [kl̥] 'I bind'	*ungl* [gl] 'I mark with fingernail'
(b)	*arregl* [kl̥] 'I sort out'	
	lladr [tr̥] 'I bark'	
	alegr [kr̥] 'I cheer'	
	obr [pr̥] 'I open'	
	fibl [pl̥] 'I sting'	

When no nasal precedes, as in (7b), the voicing contrast is not observed, and coda obstruents are voiceless. In spectrographic measurements Dols observed markedly shorter stop durations for the underlying voiced stops of (7a) in comparison with the underlyingly voiceless stops in the first column. These differences were largely confirmed experimentally by Recasens & Pallarès (2000), who, however, observed no voicing in the final liquids in any of the sets. Unlike Llach (1998), they observed no significant difference in the duration of the nasal segment in words of type (7a). One of their two speakers maintained a voicing contrast (voicing during stop closure) in just one of the clusters in (7b), /Vbl/ (lexical item not stated).

The difference observed in examples like those in (7a) was also confirmed in the experimental study performed by Llach (1998). She found that in the examples with underlying voiced stops the duration of the nasal was significantly longer, and the duration of the stop closure was significantly shorter, than in corresponding examples with voiceless underlying stops.[3] As if to compensate, the length of the stop closure was shorter in the clusters with underlying voice, with the consequence that the overall duration of the nasal+stop sequences in each type was approximately equal. The precise proportion of nasal and stop duration is unclear from Llach's text; but, recalculating from her graphs, it seems that in the nasal+voiceless stop sequences the nasal duration accounted for some

[3] Preceding vowel length was significantly longer only before /ndɾ/.

39 per cent of the whole, while in the nasal+voiced stop sequences the nasal accounted for some 56 per cent of the whole. She also examined final stop+liquid clusters without a preceding nasal, using examples like those in (7b) along with examples containing underlying voiceless stops, such as *estupr* 'I rape', *xupl* 'I suck', *perpetr* 'I perpetrate'. Here greater vowel duration was observed before underlying voiced stops than before underlying voiceless ones (about 42 per cent longer, to judge from the graphs); in this case, however, only in the /pl/ cluster was the stop duration itself significantly greater among the underlying voiceless stops than among the underlying voiced ones. Llach observed generally, in all the types of final stop+liquid cluster, the presence of audible stop bursts, such as are often absent in strictly word-final stops. Interestingly, the differences in duration of the preceding segment, whether nasal (as 7a) or vowel (as 7b), before underlying voiced or voiceless stops were of the same order. Yet native speakers had not previously remarked on the presence of a contrast in final stop+liquid clusters not preceded by a nasal, despite strong morphophonemic (and orthographic) support. This may suggest that it is only the difference in stop duration that provides a cue which provokes conscious awareness in some hearers of an underlying voicing contrast in the case of nasal+stop+liquid clusters, not the duration of the preceding segment. Or it may be that nasal length is itself more perceptible than vowel length. To test this one would need to investigate the realization of underlying contrasts of /mb/, /nd/, and /ng/ with /mp/, /nt/, and /nk/ respectively, in word-final position. These potential contrasts, which are not restricted to Balearic dialects, are always reported as subject to the same voicing neutralization as other obstruent pairs. The general issue raised by 'incomplete neutralization' is, of course, the possible mismatch between production and perception, an issue which is confronted more directly in approaches like that of Boersma (1998; 1999; 2000), who argues for distinct grammars of production and perception.

5.3 ACCOUNTING FOR VOICING NEUTRALIZATION AND ASSIMILATION

In recent phonological work two approaches to neutralization have been developed, which, following Steriade (1997), can be identified as *licensing by prosody* and *licensing by cue* respectively. The licensing by prosody approach is developed by Beckman (1998) and Lombardi (1999; 2001), and applied specifically to Catalan voicing neutralization by Beckman (1998: §1.3) and by Jiménez (1999: 165-88). The central element of this approach is the universal difference between onsets and codas when it comes to realizing consonantal contrasts. Contrasts that may be realized in onsets may be neutralized in codas. From an OT perspective, this means that Faithfulness (to underlying featural specifications) in onsets outranks Markedness constraints which penalize the realization of features, either generally or in specific contexts such as codas. Beckman's summary (§1.3.3) of constraint interactions governing Catalan obstruents is given here in (8).

(8) Constraint interactions governing Catalan obstruents (Beckman 1998: §1.3.3)

Ranking	Result
*VOICEDOBSTR	Free-standing coda obstruents must be voiceless.
IDENT-ONSET(voice) » *VOICED OBSTR	Onset obstruents may be voiced or voiceless.
AGREE(voice) » *VOICEDOBSTR	Clusters agree in voicing, even if voiced obstruents are derived from underlying voiceless segments.
AGREE(voice) » IDENT(voice)	Clusters agree in voicing, even if deviations from the underlying [voice] specifications are required.
IDENT-ONSET(voice) » IDENT(voice)	When unfaithfulness is compelled, coda obstruents, rather than onsets, will be unfaithful.

While the constraint interactions listed in (8) give the general flavour of the licensing-by-prosody approach, they are not actually adequate to generate the neutralization data set out in (4–7) above. For one thing, Beckman's AGREE (voice) abbreviates: 'Obstruents in a cluster must agree in voicing [sc. with each other]' (1998: §1.3.1.1), so does not account for voicing assimilation before sonorant consonants (as in (4) and (5)). She does not consider the issue of the voiced realization of word-final prevocalic sibilants as in (6), or non-neutralized Balearic final clusters as in (7), neither of which patterns is consistent with the details of the account she proposes.[4]

The licensing-by-cue approach is developed by Steriade (1997; 1999; 2001b), and is applied to Catalan voicing neutralization and assimilation by Llach (1998). This approach aims to base constraints licensing specific features not just on hierarchies derived from cross-linguistic typologies, but more specifically on differences in perceptual cues to be found in different phonetic contexts. Steriade convincingly shows that although, for many features, onsets provide better cues than codas (because many consonant contrasts such as place or voicing are strongly signalled in consonant–vowel transitions, or more generally in consonant–sonorant transitions), other features, for example retroflexion or pre-aspiration, are preferentially signalled in vowel–consonant transitions, i.e.

[4] While Beckman's AGREE(voice) constraint is inadequate because it is too narrow, Jiménez's version (1999) $\text{PROPAGUEU}_{\text{Laringi}}$ ('$\text{SPREAD}_{\text{Laryngeal}}$'), though appropriately ranked below IDENT-ONSET(voice), is too general, since it does not rule out spreading voice from an adjacent vowel. If one follows Jiménez (166) in taking Laryngeal to be a privative feature, his constraint ranking IDENT-ONSET(voice) » $\text{SPREAD}_{\text{Laryngeal}}$ » *VOICEDOBSTR » IDENT (voice) ought, for *actual* 'present', to prefer *[ag.tu.ˈal], in which voicing is spread to the velar stop from an adjacent vowel, to the correct [ak.tu.ˈal], which appears to show a violation of the $\text{SPREAD}_{\text{Laryngeal}}$ constraint. If Laryngeal were not taken as privative, then *[ag.tu.ˈal] and [ak.tu.ˈal] would each incur an equivalent violation of $\text{SPREAD}_{\text{Laryngeal}}$, as in this case *[ag.tu.ˈal] would fail to show spreading of voicelessness to [g] from the adjacent [t]; evaluation would pass to *VOICEDOBSTR, by which criterion [ak.tu.ˈal] would win.

typically in codas. Steriade (1997: 6) identifies seven phonetic properties known to be available as cues to voicing contrast in obstruent stops:

- closure voicing,
- closure duration,
- duration of preceding sonorant,
- F_1 values in preceding sonorant,
- burst duration and amplitude,
- voicing onset time (VOT), and
- F_0 and F_1 values at onset of voicing in a following sonorant.

Table (9) (after Steriade 1997: 6) indicates which of these features are available for obstruent stops in a series of contexts.

(9)

	Internal cues		Offset cues		Onset cues		
Context	Closure voicing	Closure duration	Son_1 duration	F_1 values in Son_1	Burst duration/ amplitude	VOT	F_0 and F_1 transitions to Son_2
Son_1_Son_2	✓	✓	✓	✓	✓	✓	✓
Obs_Son_2	✓	✓			✓	✓	✓
#_Son_2	✓				✓	✓	✓
Son_1_#	✓	✓	✓	✓	(✓)		
Son_1_Obs	✓	✓	✓	✓			
Obs_Obs	✓	✓					
Obs_#	✓	✓					
#_Obs	✓						

Son_1 = preceding sonorant; Son_2 = following sonorant; Obs = obstruent.

Although as a general rule, more cues are better than fewer cues, Steriade remarks (1997: 9) that the cues should also be weighted, inasmuch as evidence suggests that onset cues have primacy over offset cues, and also over a combination of offset and internal cues. This 'weighted cue' approach is what is supported by the cross-linguistic implicational hierarchy of contexts for obstruent stop voicing contrasts reported by Steriade (1997: 8) which runs as in (10).

(10) #_Obs, Obs_# ⊃ Son_Obs ⊃ Son_# ⊃ #_Son ⊃ Son_Son
e.g. bsa vs. psa, asb vs. asp; absa vs. apsa; ab vs. ap; ba vs. pa; aba vs. apa

In Catalan, VOT is known to be a major cue to voicing contrasts in stops. Recasens (1986: 74) reports VOT for voiceless plosives of between +15 and +58 ms, and of voiced stops of between −65 and −98 ms. As regards different classes of following sonorants, Llach observes in her account of stop onset cues that following nasals obscure the cues without masking them altogether, while

following liquids behave much the same as vowels do in this respect. In particular, she notes (1998: 27–8) that opening the velum lowers the oral pressure and reduces the stop burst, while the low intensity of the nasal segment and the presence of antiformants may mask the remaining onset cues. That is to say, we should expect the following implication for obstruent stop voicing contrasts between the different _Son contexts: _Nasal ⊃ _Liquid, _Vowel. Steriade (1997: 30-36) also investigates differences between vowels and sonorant consonants in the facilitation of cues to voicing in a preceding obstruent (though her investigation does not make a distinction between nasals and liquids). Based on comparison of various Polish and Russian sequences, she concludes that a perceptibility scale can be established distinguishing three contexts: (a) a following V, RV, or syllabic R (where R = sonorant consonant), (b) a following R# (word-final sonorant), and (c) a following non-syllabic R and another obstruent (RO). Neither Russian nor Polish neutralizes voice in context (a); in context (b) Polish neutralizes but Russian does not; in context (c), both Russian and Polish neutralize.[5] Balearic Catalan (see (7)) would resemble Polish in this regard, as far as perception is concerned, showing neutralization in VCɾ# and VCl# clusters. However, in Catalan a preceding nasal is sufficient to preserve contrast, in VNCɾ# and VNCl# clusters; that is to say, the only context in which word-final voicing contrast is preserved in Catalan is one in which an obstruent is both preceded by a nasal and followed by a liquid.

As far as fricatives specifically are concerned, Llach mentions three potential cues to voicing contrasts:

- friction duration,
- voicing during friction, and
- transitions to the following segment.

The transition cues for voice in fricatives are also likely to be masked by a following nasal, and perhaps by a liquid as well, or rather, specifically by a liquid after a homorganic fricative. Neutralization + voice assimilation seems entrenched for [Sl], [Sr], [ʃl], [ʃr] in a different way than with [fl], [fr], where the fricative is clearly voiceless, and [vl], [vr] seem more like accidental gaps, as mentioned above (§5.1). It may be homorganicity that contributes to the absence of a /tl/ – /dl/ contrast, though here the parallel /tɾ/ – /dɾ/ contrast seems solid.

A final issue to investigate from the licensing-by-cue perspective is that of the direction of voicing assimilation. If we consider just obstruent sequences for the moment, we will want to express the fact that sequences dissimilar in voicing, such as [kz] or [gs], are more marked than sequences of the same voicing: [ks], [gz]. This markedness will follow from the implicational observation: if a language has obstruent clusters disagreeing in voicing, then it also has clusters

[5] The appendix to this chapter outlines Steriade's 'P-map' approach to the perceptual confusability of voicing cues in various contexts.

agreeing in voicing, but not conversely. So we have a universal markedness constraint hierarchy (11).

(11) * [αvoice, −son] [−αvoice, −son] » * [αvoice, −son][αvoice, −son]

Given (11), I further need to explain why regressive voicing assimilation is less marked than progressive assimilation. This, in fact, follows from the greater strength of onset voicing cues than offset voicing cues (see (9) above, and Steriade 2001b: 35–6; 1997: 43–4). Deriving the hierarchy again from the cues, the following ranking (12) is seen, in which faith to pre-sonorant voice outranks faith to pre-obstruent voice.

(12) IDENT [±voice]/_[+son] » IDENT [±voice]/ _[–son]

Given inputs /kzV/ and /gsV/, the markedness constraint *[αvoice, −son] [−αvoice, −son], together with the ranking in (12), will declare the respective outputs [gzV] and [ksV] superior to the progressively assimilated [ksV] and [gzV], as is appropriate.

Among the examples given above in (5) and (6) is seen evidence of 'post-lexical' effects. These effects become evident in particular when a word-final obstruent is followed by a word-initial vowel (including glide) or liquid. Within words, as seen in (1) and (2), voicing contrast is available for obstruents before vowels, glides, or liquids ([+approximant]), but word-final obstruents maintain in corresponding contexts the neutralization found in the isolation or utterance-final contexts. (Before other consonants, word-final obstruents neutralize and assimilate voice (5) just as they do internally (4).) Recall that before syllabic vowels, word-final stops and /f/ are voiceless (*sap ajudar* [p] 'knows how to help'), as in utterance-final position. Before approximants, word-final stops are voiced (*sap riure* [b.r] 'knows how to laugh'; *cap iogurt* [b.j] 'no yoghurt'). In contrast, before syllabic vowels, word-final sibilants are voiced (*mateix element* [.ʒ] 'same element', *llargs anys* [g.z] 'long years'); they are also voiced before approximants (*saps limitar* [bz.l] 'know.2SG.PRS.IND how to limit'). Combinations of stems and inflexional affixes behave just like stem-internal cases: there is no voicing neutralization. The same is true of combinations of stems and derivational suffixes—virtually all derivational suffixes are vowel-initial. This general pattern of maintaining voicing contrast is illustrated in the right-hand column of (3). Combinations of pronominal clitic and verb also reflect the word-internal pattern, with contrast maintained, as illustrated in (13) (Recasens: 1993: 171–2; Palmada 2002: 258–9).

(13) Verb-clitic combinations with voice contrasts at contiguous edges

Input voiced	Input voiceless
ens estimem [n.z] 'we love each other'	*s'estimen* [s-] 'they love each other'
reb-ho [ˈrɛβu] 'receive.IMP-it.NEUT'	*tus-hi* [ˈtusi] 'cough.IMP-there'

reb-la [ˈrɛb.lə] 'receive.IMP-it.F'
perd-ho [ˈpɛrðu] 'lose.IMP-it.N'
fuig-hi [ˈfuʒi] 'flee.IMP-there'
cus-ho [ˈkuzu] 'sew.IMP-it.N'
aplaudeix-ho [əpləwˈðɛʃu] 'applaud.IMP-it.N'
escup-ho [əsˈkupu] 'spit.IMP-it.N'

With respect to prefix–stem combinations, the evidence concerning maintenance of voicing contrast is limited and not without ambiguity. The prefixes *ex-* and *des-* show final [z] (voiced) before vowel-initial stems (e.g. *exalçar* [əgzəlˈsa] 'to exalt', *desanimar* [dəzəniˈma] 'to discourage') but there is no reason not to think the [z] is underlying, i.e. there is no positive evidence for neutralization here. The prefixes *ab-*, *ob-*, and *sub-* all display behaviour distinct from word-internal sequences. Before roots beginning with liquids the final labial in these prefixes remains in a coda, i.e. it does not resyllabify to make a complex onset with the following liquid as in (14).

(14) (a)

ab-	*ob-*	*sub-*
abreacció [b.r] 'abreaction'	*obrepció* [b.r] 'obreption'	*subratllar* [b.r] 'to underline'
abrogar [b.r] 'to abrogate'	*obliterar* [b.l] 'to obliterate'	*subrepció* [b.r] 'subreption'
ablactació [b.l] 'weaning'		*subrogar* [b.r] 'to substitute'
		subliminar [b.l] 'subliminal'
		sublitoral [b.l] 'sublittoral'
		sublingual [b.l] 'sublingual'

(b)

abrupció [.βr] 'abruption'	
ablació [.βl] 'ablation'	*oblació* [.βl] 'oblation'
ablució [.βl] 'ablution'	*oblong* [.βl] 'oblong'

The syllabification illustrated in (14a), which differs from that found in simplex words such as *abric* [əˈβrik] 'overcoat', *oblic* [uˈβlik] 'oblique', or *sublim* [suˈβlim] 'sublime', reflects awareness of the morphological structure involving a prefix+root combination. In (14b) I add some examples in which morphological complexity might be argued for, but whose pronunciation does not apparently reflect it (and Bruguera 1990 records [βl] also in *ablactació*). When the prefix *sub-* occurs before a vowel-initial root, it is pronounced [sup], as in *subaltern* 'substitute', *subatòmic* 'sub-atomic', *subocupació* 'under-employment', *suburbà* 'suburban', and several other words.[6] *Ab-* in *aborigen* 'aboriginal' (cf. *origen* 'origin'), *aberrant* 'aberrant' (cf. *errant* 'errant'), *abús*

[6] *Subordinar* 'to subordinate' and *suburbi* 'suburb' have [β], a pronunciation which implies that the presence of prefix *sub-* is not recognized.

'abuse' (cf. *ús* 'use'), *abintestat* 'intestate' (cf. *intestat* 'intestate') is pronounced with [β], though an informant also accepts [p] in *abintestat*. There seem to be no examples of the prefix *ob-* before a vowel-initial root. The difference between *sub-* with predominant [p] in vowel-initial contexts and *ab-* with predominant [β] in vowel-initial contexts seems to offer some slight evidence for an underlying contrast, between /p/ in *sub-* and /b/ in *ab-*, in which case the realization of both with [b] before root-initial liquids as in (14a) would represent normal coda neutralization, just as we find with word-final obstruents. Alternatively, only *sub-* would display regular sandhi behaviour, while most *ab-* words would be treated as if they were morphologically simple. The pronunciation of *sub-* is problematic for the licensing-by-cue approach, as pointed out by Wetzels & Mascaró (2001: 225–6, fn. 20), and as acknowledged by Steriade (1997: fn. 26). Assimilation in voicing of /p/ to following /l/ or /r/, rather than resyllabification with them as [.pl], [.pɾ], cannot plausibly be accounted for as a 'cyclic', or paradigm uniformity, or output–output correspondence effect, as there is no independent word form *[sup] (or *[sub]). Neutralization of prefix-final /p/ and /b/ before word-initial liquids would make the problem worse for the licensing-by-cue approach, since from that perspective, if a following liquid licenses a voice contrast, it should do so independently of whether the obstruent is in an onset or a coda. The account I favour for these prefixes makes their syllabification subject to an anchor constraint ANCHOR(Stem, σ, L): 'Any element at the left edge of a Stem has a correspondent at the same edge of a syllable' (McCarthy 2003b: 89). Like UNIFORMσ (§3.1), ANCHOR(Stem, σ, L) is dominated by ONSET; its effect becomes apparent only when a syllable already has an onset.

5.4 TOWARDS AN OT ACCOUNT OF VOICING NEUTRALIZATION IN CATALAN

The licensing-by-cue approach is in principle an attractive one, since it seeks to explain the differential distribution of phonetic properties—in this case voicing—directly in terms of different phonetic environments, considering their articulatory and perceptual characteristics. It is properties of these environments that favour or disfavour the realization of voicing contrasts, or the direction and extent of voicing assimilation. Licensing by prosody appeals rather to elements of prosodic organization—specifically, syllable structure—which is held to be psychologically real, but which is only indirectly manifest in the stream of speech. Of course, there is good evidence that syllable structure is itself organized on the basis of contrast cues. However, syllable organization is one step removed from the data that motivate it. It would not be surprising, then, if prosodic structure were organized on the basis of generalizations from contextual cues, but came to have properties that were partly independent of the phonetic facts that originally motivated them. Comparing the two approaches, Steriade asks (1997: 50), 'What would count as genuine evidence for syllable-final devoicing?' She goes on: 'The

simple answer is: any system that allows us to compare voicing maintenance in onset O[bstruent]R [=sonorant] sequences with voicing neutralization in heterosyllabic O.R. Thus the hypothetical language . . . distinguishes voiced obstruents in the OR sequences functioning as onsets, but neutralizes voicing in every other obstruent-C sequence, including in heterosyllabic O[.]R.' She mentions Catalan in a footnote at this point as such a possible language, but does not pursue the issue in the absence (to her) of sufficient data. It will be seen that Catalan does indeed illustrate the syllable-final devoicing pattern (strictly, syllable-final voicing neutralization) whose existence she seeks to cast doubt on.

I have already remarked on one aspect of Catalan voicing distribution that refutes a pure licensing-by-prosody approach: the Balearic word-final cluster contrast illustrated in (7a), where *concentr* [-ntɾ] 'I concentrate' ends in a cluster distinct from *cilindr* [-ndɾ] 'I roll'. These contrasts require a highly ranked faithfulness constraint (15) making no reference to syllable structure, as licensing by cue requires.

(15) IDENT[±voice]/[+nasal]___[+cons, +approx] (IDENTVceNCR)

The constraint IDENTVceNCR is understood to derive from the phonetic and perceptual hierarchies mentioned in (9) and (10) and the discussion surrounding them. It necessarily outranks more general IDENTVce constraints, such as IDENT[±voice]/__[+cons, +approx].

However, there are further facts of Catalan that indicate that one cannot rely on a cues-based approach, to the exclusion of a prosodic approach. The examples in (4d) are relevant here. In particular, there are no voicing contrasts /sr/ – /zr/, /sl/– /zl/, or /tl/ – /dl/; only the voiced obstruent occurs. Other obstruents do contrast before liquids (unless neither the voiceless nor the voiced obstruent can occur there—alveolo-palatals do not occur at all before liquids). The licensing-by-cue approach prompts us to look for characteristics of [s] to explain why it might not be adequately distinguished from [z] before liquids, or characteristics of [t] to explain why it might not be adequately distinguished from [d] before [l]. But this seems entirely to miss the point. The segment sequences admitting obstruent voicing contrasts in Catalan are all and only those that can occur in onsets. The sequences /sr/ – /zr/, /sl/ – /zl/ do not occur in onsets inasmuch as no cluster of a sibilant and a consonant appears in a Catalan onset. Of course, it is legitimate to seek an explanation of why such onset sequences are ill-formed—the issues are discussed here in §3.1. But there is no reason at all to think it has anything to do with the realizability of voicing contrasts. In §3.1 I also consider why there are no /tl/ or /dl/ ([dl] ∼ [ðl]) onsets in Catalan, but voicing contrast is not plausibly involved. The relevant fact about voicing contrast in these clusters, then, is that they are all syllabified as coda+onset, and, with the exception of the Balearic -VNCR# clusters mentioned above, Catalan does not allow contrastive voice in codas. A plausible reason for not allowing contrastive voice in codas is that *in general* coda environments offer poor cues (or weaker cues) to voicing, since codas necessarily lack most or all of the privileged onset cues

listed in (9). Languages can quite naturally generalize to prosodic contexts as a whole, abstracting from the more specific phonetic facts which directly license contrasts.

A more subtle problem for the cue approach arises when one considers a word sequence like *poc realista* /ˈpɔk reaˈlista/ [ˈpɔg.reəˈlistə] 'not very realistic'. In order to focus on the problem, I assume that the constraints ONSET and UNIFORMσ (see §3.1) are highly ranked, as is a paradigm uniformity constraint enforcing correspondence between [continuant] values in words and phrases.[7] Thus candidates like *[ˈpɔɣ.ɾeəˈlistə] and *[ˈpɔ.gɾeəˈlistə] are ignored. Unassimilated candidates like *[ˈpɔk.reəˈlistə] are also ignored, assuming a high-ranking voicing assimilation constraint (for which, see below). The constraints to focus on here are a correspondence constraint enforcing corresponding voicing between words and phrases, IDENTWd-Phr[±voice] and the faithfulness constraint preserving underlying voicing contrasts in 'adequately rich' cue environments: IDENTI-O[±voice]/___([+approx])V. I also include markedness constraints penalizing voiceless sonorants and voiced obstruents.

(16)

	/ˈpɔk#reaˈlista/ [ˈpɔk] [reəˈlistə]	IDENTWd-Phr[±voice]	IDENTI-O[±voice] /__([+approx])V	*[–vce, +son]	*[+vce, son]
(a)	☞ ˈpɔg.reəˈlistə	*	*		*
(b)	ˈpɔg.r̥eəˈlistə	*	*	*!	

In tableau (16) candidate (16a) has one violation of IDENTWd-Phr[±voice] inasmuch as the word [ˈpɔk] is realized as [ˈpɔg]. Candidate (16b) has an equivalent violation in that [reəˈlistə] is realized with an initial voiceless [r̥]. The same candidates violate input–output faithfulness in the same way, since in the word *poc* /k/ is underlying. Only the ranking *[−vce, +son] » *[+vce, −son] ('voiceless sonorants are more marked than voiced obstruents') declares (16a) the winner. However, this ranking produces the wrong result for a word like Balearic *lladr* /ˈʎadr/ [ˈʎatr̥] 'I bark' (see (7b)).

(17)

	/ˈʎadr/	IDENTWd-Phr[±voice]	IDENTI-O[±voice] /__([+approx])V	*[–vce, +son]	*[+vce, –son]
(a)	ˈʎadɾ				*
(b)	💣 ˈʎatr̥			*!	

[7] Or rather, with the same effect, a high-ranked constraint distributing stop and fricative allophones of non-strident obstruents, for which see §10.1. This allophony constraint seems to be undominated in Catalan.

In (17) IDENTWd-Phr[±voice] is irrelevant, since only single-word candidates are compared. Word-final -Cr# does not match the 'rich voicing cues' environment, so IDENTI-O[±voice] /__([+approx])V cannot choose between them. Then *[−vce, +son] eliminates the candidate that should win. It is conceivable that this problem can be fixed up in the cue approach, but the fact remains that provided one adopts a positional faithfulness version of IDENTI-O[±voice]/__([+approx])V, namely IDENTI-O[±voice]$_{\text{Onset}}$, the problem in (16) and (17) disappears. Onset faithfulness prefers [ˈpɔg.reəˈlistə] to [ˈpɔk.r̥eəˈlistə] and *[−vce, +son] can be eliminated from the picture, leaving *[+vce, −son] to prefer [ˈʎatr̥] over [ˈʎadɾ]. Or, to be more explicit, *[−vce, +son] is demoted below a constraint disfavouring a voicing contour in a margin. Given the contexts needed for effective expression of voicing (see (9)), a [+voice][−voice] cluster is highly marked in a word-initial onset, and similarly a [−voice][+voice] cluster is highly marked in a word-final coda. In each case, in addition to the effort involved in switching voicing within a cluster, the voiced segment is compelled to rely on internal cues to voicing. I call this constraint VOICEGRADIENT (VOGRAD) (18).

(18) VOICEGRADIENT (VOGRAD): Within a margin, a voiced segment is not further from the nucleus than a voiceless segment. (After Greenberg 1978.)

The tableau for *lladr* now looks like (19).

(19)

	/ˈʎadr/	IDENTI-O[±voice]$_{\text{Onset}}$	VOGRAD	*[+vce, –son]	*[–vce, +son]
(a)	ˈʎadɾ			*!	
(b)	ˈʎatɾ		*!		
(c)	☞ ˈʎatr̥				*

VOGRAD » *[–vce, +son]

And the preferred tableau for *poc realista* is now as in (20).

(20)

	/ˈpɔk#reaˈlista/ [ˈpɔk] + [reəˈlistə]	IDENTI-O [±voice]$_{\text{Onset}}$	VOGRAD	*[+vce, –son]	*[–vce, +son]
(a)	☞ ˈpɔg.reəˈlistə			*	
(b)	ˈpɔg.r̥eəˈlistə	*!			*

IDENTI-O[±voice]$_{\text{Onset}}$ » *[+vce, –son]

Now it is possible to proceed to consider which constraints outrank those illustrated in (19) and (20). I return first to the correspondence constraint mentioned before (16): IDENTWd-Phr[±voice], and defined here in (21).

(21) IDENTWd-Phr[±voice]: The value for [±voice] in a prosodic word is the same as that of its correspondent in a phrase.

This constraint can be seen to outrank the I-O faithfulness constraint IDENTI-O [±voice]$_{\text{Onset}}$ in the case of a word-final neutralized stop before a vowel-initial word, as in the example from (6) *sap ajudar* / ˈsab#aʒuˈdaR/ [ˈsa.pəʒuˈða] 'knows how to help'. Though the prevocalic, onset position of the root-final /b/ is one that could, in principle, permit the realization of underlying voicing, it is in fact realized as a voiceless stop, as corresponds to the citation or utterance-final form *sap* [ˈsap]. The relevant candidates are compared in tableau (22).

(22)

	/ˈsab#aʒu' dar/ [ˈsap] + [əʒuˈða]	IDENTWd-Phr[±voice]	IDENTI-O[±voice]$_{\text{Onset}}$
(a)	☞ ˈsa.p əʒuˈða		*
(b)	ˈsa.b əʒuˈða	*!	

IDENTWd-Phr[±voice] » IDENTI-O [±voice]$_{\text{Onset}}$

However, IDENTWd-Phr[±voice] must in turn be outranked by the coda-voicing assimilation constraint, because for *sap riure* 'knows how to laugh', [ˈsab.ˈriwɾə] is better than *[ˈsap.ˈriwɾə]. It would be attractive to adopt the licensing-by-cue perspective on voicing assimilation, using constraints of the form of Steriade's (2001b: 33) such as *[αvoice, −son][−αvoice, −son]: 'Adjacent obstruents agree in voice.' The direction of assimilation would follow from constraints enforcing the underlying voice of the second obstruent, in a pre-sonorant position that licenses contrast preservation, or before an unmarked word-final voiceless consonant. In Catalan, however, an obstruent agrees with a following segment in voicing not only before another obstruent, but also before a nasal, before a liquid and indeed before a glide, *provided the obstruent is in a coda*, as the examples in (4) and (5) show. By contrast, no obstruent that is in an onset has to agree in voicing with a following segment. Thus, as far as voicing agreement is concerned, the context of the constraint relevant to Catalan is ineluctably prosodic. It is also undominated; i.e. there are no surface exceptions. The constraint is formulated in (23); I adopt Beckman's label AGREE(voice) (1998: §1.3.3).

(23) AGREE(voice): *[−son, αvoice]$_{\text{Coda}}$[−αvoice] (A coda obstruent and a following segment must agree in voicing.)[8]

The role of AGREE(voice) is illustrated in tableau (24) with the string *sap riure* [ˈsab.ˈriwɾə] 'knows how to laugh'.

[8] The representation here with]$_{\text{Coda}}$ simplifies for representational reasons prosodic tree dominance: provided the relevant segment is in a coda, it is immaterial whether the following segment is in the same coda or in an onset. Example (27b) below shows *pots* 'can.2SG' [pɔts#] ~ [pɔd.z#V].

(24)

/'sab#'riwɾe/ ['sap] + ['riwɾə]	AGREE(voice)	IDENTWd-Phr[±voice]	IDENTI-O [±voice]$_{Onset}$	*[+vce, −son]
☞ 'sab.'riwɾə		*		*
'sap.'riwɾə	*!			

AGREE(voice) » IDENTWd-Phr[±voice]

5.5 WORD-FINAL SIBILANTS

Example set (6) also illustrates the fact that word-final sibilants, neutralized for voice contrast, are pronounced voiced before a vowel-initial word (the sibilant is resyllabified into the onset, as discussed in §3.1.1). The data are repeated in (25).

(25) *cas extrem* [z] 'extreme case' *cops amagats* [bz] 'hidden blows'
mateix element [ʒ] 'same element' *pots ajudar* [dz] 'may.2SG help'
mig any [d͡ʒ] 'half a year' *llargs anys* [gz] 'long years'

The phonological reason for this voiced output of word-final prevocalic sibilants is not obvious. The output violates the correspondence constraint IDENTWd-Phr[±voice], without being driven by any general assimilation process. Voiced and voiceless sibilants are lexically in contrast in prevocalic position. The output also violates the general markedness of voiced obstruents. Llach (1998: 65) suggests that the function of the process is to assist in segmentation and word recognition: a prevocalic voiceless sibilant cannot be word-final. But such an effect follows from any word-edge neutralization, not specifically of voicing. Hence, it is also the case that, in Catalan, a prevocalic voiced non-sibilant obstruent cannot be word-final. So it still needs to be explained why the neutralized prevocalic sibilants are voiced, specifically.

It may rather be that the phenomenon does not have a good synchronic functional explanation. Note that voiced sibilant 'liaison' is shared by Portuguese, Catalan, Occitan, and French (at least) among Romance languages; cf. French *six* [sis] 'six', *six heures* [si.zœʁ] 'six o'clock'. (In French the phenomenon is extended to the labio-dentals: *neuf* [nøf] 'nine', *neuf heures* [nø.vœʁ] 'nine o'clock'.) All these languages experienced a lenition sound change, affecting non-initial single intervocalic obstruents. The process became opaque when remaining voiceless geminate obstruents, including /-ss-/, standing in original intervocalic voicing environments, were simplified. Thus it may be argued that voiced sibilant liaison is synchronically an unnatural morphophonemic fact of the grammars of certain languages,[9] and no longer a natural process in a natural environment. I here call the constraint enforcing voiced sibilant liaison

[9] There is little reason to doubt that voiced sibilant liaison was also once part of Spanish grammar, which was lost when all voiced sibilants became voiceless in the 15th and 16th centuries. In the same way, voiced sibilant liaison has been lost in the *apitjat* dialects of Catalan.

LAZYSIBILANTS (LAZYSIB) (26), in recognition of the origin of the process in lenition. (LAZY constraints in current Catalan lenition are discussed in §10.1.)

(26) LAZYSIBILANTS (LAZYSIB): Word-final sibilants are voiced before a vowel.

Tableau (27) illustrates the effect of LAZYSIB in the phrases *mateix element* 'same element', *pots ajudar* 'may.2SG help'

(27)

	/maˈteʃ#eleˈment/ [məˈteʃ] + [ələˈmen]	LAZYSIB	AGREE (voice)	IDENTWd-Phr [±voice]	IDENTI-O [±voice]$_{Onset}$	*[+vce, −son]
(a)	☞ məˈte.ʒ ələˈmen			*	*	*
	məˈte.ʃ ələˈmen	*!				
(b)	/ˈpɔd+z#aʒuˈdar/ [ˈpɔts] + [əʒuˈða]					
	☞ ˈpɔd.z əʒuˈða			*		***
	ˈpɔt.s əʒuˈða	*!			*	*
	ˈpɔt.z əʒuˈða		*!	*		**
	ˈpɔd.s əʒuˈða	*!	*	*	*	**

At the end of §5.1 I mentioned Catalan dialects with rather divergent consonant sandhi, in which word-final non-strident obstruents before vowels are realized not as voiceless stops, but as voiced fricatives, as in (28).

(28) *Alacantí*
sap ajudar [ˈsaβ ad͡ʒuˈðaɾ] 'knows how to help'
pot ajudar [ˈpɔð ad͡ʒuˈðaɾ] 'may.3SG help'
llarg any [ˈʎaɾɣ ˈaɲ] 'long year'

In these varieties it appears that a constraint like LAZYSIB applies to all obstruents, not just to sibilants, with the result that the effect of the correspondence constraint IDENTWd-Phr[±voice] is altogether obscured. In word-final position voiceless obstruents are found only before a voiceless sound or utterance-finally. The fricative realization of the word-final voiced non-strident obstruents in these varieties follows from the general lenition constraint distributing voiced stops and fricatives in Catalan (§10.1): fricatives in onsets between continuants, stops elsewhere. In fact, in *alacantí* the distribution of voiced fricatives is yet more general; all post-continuant voiced obstruents are fricatives, whether in onsets or in codas. Where most varieties have, for *sac gros* 'big bag' [ˈsag.ˈɡɾɔs], *alacantí* has [ˈsaɣ.ˈɣɾɔs], or, with simplification of geminates, [ˈsa.ˈɣɾɔs]. Ignoring degemination, the constraints particular to these varieties (*alacantí* etc.) are spelt out in (29).

(29) LAZYVOICE$_{alacanti}$ (LAZYVO$_{al}$): Word-final obstruents are voiced before a vowel.
LAZYCONTINUANT$_{alacantí}$ (LAZYCONT$_{al}$): Non-strident voiced obstruents spread [+cont] from a preceding segment. (*[+cont][−cont, −son, −strid, +vce])

Tableau (30) illustrates candidate evaluation for *sac gros* 'big bag' and *sac obert* ['sa.ɣ o.'βɛɾt] 'open bag' in *alacantí* (['sa.k o.'βɛɾt] in other varieties).

(30)

alacantí /'sak#'gɾɔs/ ['sak] + ['gɾɔs]	AGREE (voice)	LAZYVO$_{al}$	LAZYCONT$_{al}$
☞ 'saɣ .'ɣɾɔs			
'sak .'gɾɔs	*!		
'sag .'gɾɔs			*!
'saɣ.'gɾɔs			*!
'sag.'ɣɾɔs			*!
/'sak#ɔ'bɛɾt/ ['sak] + [o.'βɛɾt]			
☞ 'sa.ɣo.'βɛɾt			
'sa.ko.'βɛɾt		*!	
'sa.go.'βɛɾt			*!

5.6 APPENDIX: STERIADE'S 'P-MAP' APPROACH TO VOICING NEUTRALIZATION

Whereas the licensing-by-cue approach developed by Steriade in her 1997 paper considers both phonetic and typological evidence for a natural hierarchy of contexts favouring voice contrasts in obstruents, her (2001b) 'P-map' paper considers the same issues from the perspective of the *confusability* of voicing cues in the various contexts, using judgements of perceptual similarity and cases of mis-identification in experimental contexts. (Another important goal of the 2001b paper is to show why, in contexts where a voice contrast cannot be adequately realized, the 'wrong' voice is a better 'repair' than other conceivable substitutes, such as metathesis, deletion of the consonant, epenthesis of a following vowel, or nasalization of the consonant.) 'The P-map is a mental representation of the degree of distinctiveness of different contrasts in various positions' (Steriade 2001b: 14). 'If the P-map identifies the pair [p]-[b] as more

similar in the context V_] than the pair [b]-[m] in the same context, then the P-map's effect on the grammar will be to rank higher the faithfulness condition corresponding to the less confusable contrast [b]-[m], hence IDENT [±nasal]/V_] » IDENT [±voice]/V_]' (p. 4). As generally in OT, Steriade seeks to derive a universal constraint ranking from some phenomena outside grammar itself. In this case, the extralinguistic phenomena are from the psychology of perception. The part of the (hypothetical) P-map that Steriade suggests for obstruent voice contrasts displays the hierarchy of contexts in (31), ordered from most to least favourable (2001b: 18; abbreviations modified here for consistency).

(31) Hypothetical P-map voicing perceptibility context hierarchy (Steriade 2001b)
V_V > C_V > V_R > V_# > V_O > C_O
(R = sonorant, O = obstruent)

The hypothetical P-map voicing perceptibility hierarchy of (31) is not exactly the same as the hierarchy of contexts Steriade established in her earlier paper (1997: 6, 9; see (9) and (10) above). In that work the contexts V_V and V_R were treated as equivalent, and as ranking higher than C_V. She does not explicitly justify the reordering in (31). In fact, her main conclusion concerns not the hierarchy in (31) but rather the ranking of faithfulness constraints relative to just one of the contexts, the word-final post-vocalic one (/V_#), where the confusability evidence supports the ranking in (32) (adapted from Steriade 2001b: §5.2).

(32) Ranking of I–O correspondence constraints by the distinctiveness scale
LINEARITY (C_iVC_j vs. C_jVC_i), DEP (Ø vs. ə) » MAX (C vs. Ø) »
IDENT [±son]/V_# » IDENT [±voice]/V_#

Given a hypothetical input /tab/ and a markedness constraint *VOICEDOBSTR/_#, the consequence of the constraint ranking in (32) is to mark a metathesis [bat] or an epenthesis [tabə] the worst kind of violation, followed by a deletion [ta], followed by a sonority change [tam] or [taw], with the least bad non-faithful output being [tap]. That is to say, wherever *VOICEDOBSTR/_# is placed above IDENT [±voice]/V_# in the ranking, [tap] rather than the alternatives will be the preferred output. (Of course, if *VOICEDOBSTR/_# is ranked below IDENT[±voice]/V_#, the faithful candidate, [tab], will be preferred.) In other circumstances non-faithful [tabə], [ta], [tam], etc. may indeed be preferred, but this will not be because *VOICEDOBSTR/_# is active.

6

CODA PLACE AND MANNER ASSIMILATION AND NEUTRALIZATION

6.1 INTRODUCTION

Catalan is relatively rich in the range of consonants that can appear in codas, though word-medial codas are restricted in length to one consonant, or to one consonant + [s] (as in *extra* [ˈɛks.tɾə] 'extra'). The diversity of word-final codas, in particular, gives scope for extensive assimilation and neutralization of contrast before following consonants. As explained in Chapter 5, voice is never contrastive in codas. With respect to place of articulation, while homorganic clusters are found, not surprisingly, there are also a number of heterorganic clusters within morphemes, though by no means all combinations of place and manner are found (Wheeler 1975/9: 254–69; 1987; Bonet & Lloret 1998: 127–55). Place and associated properties of morpheme-internal codas are the focus of §6.2. Issues of manner assimilation—lateral and nasal—are raised in §6.3. When words are adjacent, all possible coda + onset combinations of input[1] place and manner can be found, and are subject to varying degrees of assimilation, especially of place. Issues particular to assimilation and most major place contrasts in word-final preconsonantal codas are treated in §6.4. Assimilation in coronal clusters is the focus of §6.5, where minor place assimilation is also explored. The focus of §6.6 is consonantal contacts in Majorcan Catalan, where the outputs are strikingly different from those observed in other varieties, and interestingly different from a typological perspective.

Recasens (1993: 80–198) offers a rich source of data on both internal and inter-word consonantal contacts, carefully explained, and with sophisticated interpretation from the perspective of articulatory phonetic theory, which is discussed below.[2] Table 6.1 sets out the range of possible coda + C combinations and their realizations in continental Catalan, insofar as these can be neatly represented by IPA characters. As place of articulation is a specific focus of this chapter, transcriptions are somewhat narrower in this respect than in the rest of the book.

[1] In phrases, 'inputs' to correspondence are phonological words, in their isolation (utterance-initial and utterance-final) forms.

[2] There is further detail of consonantal contact, especially with respect to dialect variation, in Recasens (1991b).

TABLE 6.1 Realization of coda-C clusters in continental Catalan (after Recasens 1993: 190, with modifications after Recasens & Pallarès 2001b). Voiced consonants only are illustrated, where both voiceless and voiced are available.

C_1 \ C_2		bil obs	bil nas	l-d fric	den obs	alv nas	alv lat	alv fric	alv trill	a-p fric	a-p nas	a-p lat	pal app	vel obs
bil obs	1 2 3	bb	bm mm	 pːf	bd	bn mn	bɫ	bz	br	bʒ	bɲ mɲ	bʎ	bj	bg
bil nas	1 2	mb	mm	 ɱf	md	mn	mɫ	mz	mr	mʒ	mɲ	mʎ	mj	mg
l-d fric	1 3	vb bb	vm mm	ff~f	vð	vn	vɫ	vz	vr	vʒ	vɲ	vʎ	vj	vɣ
den obs	1 2 3	db bb	dm mm	tf pːf	dd	dn̪ nn	dɫ̪ ɫɫ	d̪ːz	d̪r	dʲːʒ	dʲɲ ɲɲ	dʲʎ ʎʎ	dj (dʲj)	dg gg
alv nas	1 2 3	nb mb	nm mm	nf ɱf	n̪d	nn	nɫ ɫɫ	nz	nr	nʲʒ	ɲɲ	nʲʎ	nj (nʲj)	ng ŋng
alv lat	1 2	ɫβ	ɫm	ɫf	ɫ̪d	ɫn	ɫɫ	ɫz	ɫr	lʒ	lɲ	lʎ	lj	ɫɣ
alv fric	1 2 3	zβ	zm	sf	zð̪	zn	zɫ	zz~z	zɹˢ rr~r	ʒʒ~ʒ	zɲ	zʎ	zj	zɣ
alv rhot	1 3	ɾβ	rm	rf	ɾð	rn	rɫ	rz	rr~r	rʒ	rɲ	rʎ	ɾj	ɾɣ
a-p fric	1 2	ʒβ	ʒm	ʃf	ʒð̪	ʒn	zʲɫ	zʲzʲ~zʲ	zʲɹˢ	ʒʒ~ʒ	ʒɲ	ʒʎ	ʒj	ʒɣ
a-p nas	1 2	ɲb	ɲm	ɲf	ɲd	ɲn	nʲɫ	nʲz	nʲr	ɲʒ	ɲɲ	ɲʎ	ɲj	ɲg̊
a-p lat	1 2	ʎβ	ʎm	ʎf	ʎd	ʎn	lʲɫ	ʎz lʲz	lʲr	ʎʒ	ʎŋ	ʎʎ	ʎj	ʎɣ
vel obs	1 2 3	gb	gm ŋm	kf	g̊d	gn (g̊n) ŋ̊n	gɫ	gz (g̊z)	gr	g̊ʒ	gɲ ŋ̊ɲ	gʎ	g̊j	gg

Within each cell, row 1 = realizations faithful to input, row 2 = coarticulation/place assimilation, row 3 = realizations involving manner assimilation. The consonantal coronal+coronal contacts are those within the bold box. The greyed cells correspond to clusters that, for principled reasons, are not found within morphemes.

Abbreviations: bil = bilabial, l-d = labiodental, den = dental, alv = alveolar, a-p = alveolo-palatal, pal = palatal, vel = velar; obs = obstruent, fric = fricative, nas = nasal, lat = lateral, rhot = rhotic, app = approximant.

6.2 MORPHEME-INTERNAL CODA–C CLUSTERS

Semivowels (j, w) and alveolar approximants (r/ɾ, ɫ) occur freely in internal codas without restrictions as to the place of a following consonant. This wide distribution reflects the privileged role in codas of glides and alveolar approximants, as consonants of high sonority and with low coarticulatory constraints (in comparison with alveolo-palatal sonorants ɲ and ʎ, for example). Geminate coda+onset clusters ($C_i.C_i$) are acceptable in principle (see §2.1.9), but there are no geminate fricatives, and there is no evidence of underlying velar obstruent geminates, i.e. geminates involving the most marked of the Catalan non-sibilant places. For the interpretation of geminate alveolo-palatal affricates [t͡ʲːʃ] and [d͡ʲːʒ] see §2.1.2 and §2.1.3.3. Among the alveolo-palatal sonorants, geminate /ʎʎ/ is attested (e.g. *espatlla* [əsˈpaʎʎə] 'shoulder'), but */ɲɲ/ is not, and this is here assumed to be an accidental gap. A full range of homorganic nasal+C clusters is found, as illustrated in (1).

(1) Homorganic nasal clusters

Bilabial	*omplir* [umˈpli] 'to fill'	*semblar* [səmˈbla] 'to seem
Labiodental	*xamfrà* [ʃəɱˈfra] 'chamfer'	*canvi* [ˈkaɱvi] ∼ [ˈkambi] 'change'
Dental	*contar* [kun̪ˈta] 'to count'	*fondo* [ˈfon̪du] 'deep'
Alveolar	*cansar* [kənˈsa] 'to tire'	*enze* [ˈɛnzə] 'simpleton'
		honra [ˈonrə] 'honour'
		enlaire [ənˈlajɾə] 'aloft'
Alveolo-palatal	*manxa* [ˈmanʲʃə] 'bellows'	*franja* [ˈfranʲʒə] 'fringe'
Velar	*tancar* [təŋˈka] 'to close'	*tingui* [ˈtiŋgi] 'hold.3SG.PRS.SUBJ'

The range of internal clusters that involve contrast among major places of articulation in the coda consonant is somewhat restricted. In considering internal clusters I leave aside Cj (e.g. *sípia* [ˈsi.pjə] 'cuttlefish'), in which [j] has characteristics of vowel place rather than consonantal, and in any case the C is in onset, not coda position. Assuming four contrasting consonantal places, Labial, Denti-alveolar, Alveolo-palatal, and Velar, within morphemes, only the following potentially heterorganic[3] place clusters are found, with examples set out in (2) below.

Denti-alveolar–Labial (2a)
Denti-alveolar–Denti-alveolar (non-geminates) (2b)
Denti-alveolar–Alveolo-palatal (2c)
Denti-alveolar–Velar (2d)
Labial–Labial (non-geminates)[4]

[3] 'Potentially heterorganic' in that possible conflicts of minor place, dental–alveolar, and bilabial–labio-dental, are considered.

[4] Only bilabial–labio-dental clusters, in varieties with underlying voiced labio-dental, e.g. *obvi* 'obvious', *obvers* 'obverse'.

Labial–Denti-alveolar (2e)
Labial–Alveolo-Palatal (2f)
Velar–Labial (nasal only) (2g)
Velar–Denti-alveolar (2h)

Observe that there are no heterorganic clusters with Alveolo-palatal in the first slot; and when Alveolo-palatal stands in the second slot, only Denti-alveolar place (or questionably Labial, see below) is allowed in the first. In the lists that follow, coda liquids and glides are mentioned only when there is some aspect of interest concerning place realization. Where variation is indicated (with ∼) the more assimilated variant is found in a less formal style.

(2) Attested morpheme-internal clusters

(a) Denti-alveolar–Labial
adverbi [db] ∼ [bb] 'adverb'
atmosfera [dm] ∼ [mm] 'atmosphere'

(b) Denti-alveolar–Denti-alveolar (non-geminates)
Dental and alveolar place are adjusted so the sequences are in fact homorganic. What is of interest is which place, dental or alveolar, wins (§6.5).

ètnic [dn̪] ∼ [nn] 'ethnic'	*agenda* [n̪d] 'appointments diary'
atleta [dɫ̪] ∼ [ɫɫ] 'athlete'	*multa* [ɫ̪t] 'fine'
lletsó [t̪s͡] 'dandelion'	*falda* [ɫ̪d] 'lap'
setze [d̪:z͡] 'sixteen'	*fusta* [st̪] 'wood'
càntir [n̪t] 'pitcher'	*esdrúixol* [zð̞] 'proparoxytone'

The following alveolar-alveolar sequences are substantially homorganic, though /n/, /l/ are front alveolar, and /s/, /z/, /r/ are back alveolar (§2.1).

llançar [ns] 'throw'	*eslora* [zɫ] 'length (of hull)'
onze [nz] 'eleven'	*tornar* [rn] 'return.INF'
conreu [nr] 'cultivation'	*parlar* [rɫ] 'speak.INF'
balneari [ɫn] 'spa'	*farsa* [rs] 'farce'
folrar [ɫr] 'line.INF'	*esmorzar* [rz] 'breakfast'
esnob [zn] 'snob'	

(c) Denti-alveolar–Alveolo-palatal

Elx [lʃ] (toponym)	*xarxa* [rʃ] 'net'
àlgebra [lʒ] 'algebra'	*conserge* [rʒ] 'concierge'

The following affricates and nasal-sibilant clusters are substantially homorganic; see below at (24), (25), (37):

esquitxar [tʲ:ʃ͡] 'splash.INF'	*manxa* [nʲʃ] 'bellows'
calitja [dʲ:ʒ͡] 'haze'	*diumenge* [nʲʒ] 'Sunday'

(d) Denti-alveolar–Velar

adquirir [tk] ∼ [kk] 'acquire.INF'	*parc* [rk] 'park'
remolcar [łk] 'tow.INF'	*orgue* [ɾɣ] 'organ'
algú [łɣ] 'someone'	

(e) Labial–Denti-alveolar

dubte [pt] 'doubt'	*comte* [mt] 'count'
súbdit [bd] 'subject'	*ambdós* [md] 'both'
hipnosi [bn] ∼ [mn] 'hypnosis'	*solemne* [mn] 'solemn'
copsar [ps] 'catch.INF'	*premsa* [ms] 'press'
	zumzejar [mz] 'hover.INF'
	somriure [mr] 'smile.INF'

(f) Labial–Alveolo-palatal (possibly bi-morphemic; see below)

objecte [bʒ] 'object'	*abjurar* [bʒ] 'abjure.INF'

(g) Velar–Labial (nasal only)

enigma [gm] ∼ [ŋm] 'enigma'

(h) Velar–Denti-alveolar

estricte [k̟t] 'strict'	*signar* [gn] ∼ [ŋn] 'sign.INF'
maragda [g̟d] 'emerald'	*fixar* [ks] 'fix.INF'
	exacte [gz] . . . [k̟t] 'exact'

The pattern in the attested heterorganic medial clusters in (2) is generally of one of two kinds: the least marked denti-alveolar place occurs either in C_1 (2a, b, c, d) or in C_2 position (2b, e, h). Apart from these patterns involving one denti-alveolar, type (2g), Velar–Labial, is rather more marked, as is type (2f), Labial–Alveolo-palatal. There is reason to think that all of the few lexemes of type (2f) may be compounds, involving the prefix-like elements *ab-* and *ob-* (see §5.3, where the behaviour of these elements in relation to lenition of /b/ is mentioned). If the (2f) type are indeed compounds, the preservation rather than place assimilation of [b] would reflect 'word juncture', with output–output correspondence being involved, ranking higher than lexeme-internal constraints on heterorganic clusters.

In the remainder of this section I attempt to show how the restricted distribution of place contrasts in Catalan medial clusters follows from motivated markedness constraints that correspond to place hierarchies derived from typological observations. Coda place contrast and assimilation has received substantial attention within OT phonology, from Jun (1995), Padgett (1995b), Lombardi (2001), and Steriade (2001a), for example. Among other things, these authors converge on a universal markedness hierarchy for place of articulation: Dorsal > Labial > Coronal (from more marked to less marked). On the basis of a typological survey of coda place contrasts, de Lacy (2002) proposes a somewhat different way of deriving constraints, and their ranking, from a scale such as the place-markedness scale. For example, instead of one markedness constraint for each place, and a fixed ranking matching the hierarchy, namely *DORSAL » *LABIAL » *CORONAL, de Lacy argues that the true cross-linguistic typology requires 'stringent' constraints, each of which must mention the most marked

member on the scale. He refers to this (2002: 32) as the *Marked Reference Hypothesis*.[5] But the stringent constraints are themselves freely rankable. In de Lacy's approach the context-free place markedness constraints would be

*{DORSAL, LABIAL, CORONAL},
*{DORSAL, LABIAL} and
*{DORSAL}

so no constraint penalizes place without penalizing the most marked place. De Lacy abbreviates DORSAL as K, LABIAL as P, CORONAL as T, and I follow this pattern here.[6] For Catalan, of course, we need to introduce an additional Alveolo-palatal place, here abbreviated ʃ, which is also strictly CORONAL, and articulatorily also 'dorsal'; consequently, the remaining lingual places are here named DENTI-ALVEOLAR (T) and VELAR (K). The evidence from cluster behaviour within Catalan is consistent with ALVEOLO-PALATAL place being more marked than VELAR. The place hierarchy relevant to Catalan is thus as in (3), from most marked to least.

(3) ALVEOLO-PALATAL {ʃ} > VELAR {K} > LABIAL {P} > DENTI-ALVEOLAR {T}

To account for consonant clusters, in addition to context-free place markedness constraints such as *{ʃ, K, P} and *{ʃ}, and marked faithfulness constraints such as IDENT{ʃ, K, P}, IDENT{ʃ}, and so on, de Lacy argues for sets of constraints within the *HETERORGANIC family (2002: 321). His *HETERORGANIC family corresponds to the ASSIM(ilate) or AGREE constraints of other authors (including AGREEVoice here in Chapter 5). The *HETERORGANIC constraint family acknowledges the fact that some heterorganic clusters are more marked than others; for example, pk, tk are more marked than kp, tp, which are more marked than kt, pt (de Lacy 2002: 339). The full set of *HETERORGANIC constraints relevant for Catalan is given in (4), following the model of de Lacy (2002: 322).

(4) The Marked-Cluster Constraints (anti-heterorganic markedness constraints)

	*{ʃ}{ʃK}	*{ʃ}{ʃKP}	*{ʃ}{ʃKPT}
*{ʃK}{ʃ}	*{ʃK}{ʃK}	*{ʃK}{ʃKP}	*{ʃK}{ʃKPT}
*{ʃKP}{ʃ}	*{ʃKP}{ʃK}	*{ʃKP}{ʃKP}	*{ʃKP}{ʃKPT}
*{ʃKPT}{ʃ}	*{ʃKPT}{ʃK}	*{ʃKPT}{ʃKP}	*{ʃKPT}{ʃKPT}

(5) Interpretation of the anti-heterorganic constraints
*XY 'Assign a violation for every pair of adjacent segments such that
(i) the first segment has a feature f_1 from set X, and
(ii) the second segment has a feature f_2 from set Y.' (De Lacy 2002: 322)

[5] The Marked Reference Hypothesis is formulated thus by de Lacy (2002: 32): 'If a constraint C refers to a scale S, C refers to the most marked member of S. If a faithfulness constraint preserves an element P in the scale, it also preserves every more marked element.'

[6] De Lacy includes LARYNGEAL {ʔ} in his hierarchies as the least marked place. Catalan offers no evidence for laryngeal consonants. I assume *M/ʔ, h, and *P/ʔ, h are all undominated.

The strongest constraint in (4) is the one at the bottom right, *{ʃKPT}{ʃKPT}, which penalizes any heterorganic cluster whatever. The most tolerant constraints are the ones towards the top left of (4), such as *{ʃ}{ʃK} which penalizes only heterorganic clusters of most marked alveolo-palatal followed by next most marked velar place, and *{ʃK}{ʃ}, which penalizes only velar–alveolo-palatal clusters. As an example, the constraint *{ʃ}{ʃKP} assigns a violation mark to all clusters that consist of an alveolo-palatal followed by a velar or a labial, such as ɲk, ʎp, ʒβ.

The Marked-Cluster constraints (4) refer to place without consideration of the interaction of place with manner. There is in addition typological evidence of a manner hierarchy of assimilability of place: Nasals > Stops > Fricatives and non-nasal sonorants (Jun 1995: 69; de Lacy 2002: 347–8). That is, nasals are most likely to undergo place assimilation to a following consonant; if a fricative or a non-nasal sonorant undergoes place assimilation, so will a stop and a nasal of equivalent place. In principle, therefore, each constraint in (4) can be multiplied by three to distinguish manner in the first element, as illustrated in (6) taking the constraint *{ʃKPT}{ʃ} as an example.

(6) $*\{ʃKPT\}_{[\pm cont]}\{ʃ\}$ $*\{ʃKPT\}_{[-cont]}\{ʃ\}$ $*\{ʃKPT\}_{[+nas]}\{ʃ\}$

The constraints in (6) are also 'stringent'. Thus no constraint penalizes heterorganic coda place without penalizing heterorganic nasal coda place. So, if heterorganic nasal codas are admissible at all, heterorganic stop and continuant codas must be also; for example, /mt/ ⊃ /pt/ ⊃ /ft/. A constraint like *{ʃKPT}{ʃ} that does not specify manner explicitly must always be understood 'stringently' by the Marked Reference hypothesis to include nasals in the X position. (Much the same effect follows from a set of marked faithfulness constraints $\textsc{Ident}\{ʃKPT\}_{[\pm cont]}$—preserving place in all manners—$\textsc{Ident}\{ʃKPT\}_{[-nas]}$—preserving place in non-nasals—$\textsc{Ident}\{ʃKPT\}_{[+cont]}$—preserving place in continuants. The relevant IDENT constraint would dominate the relevant marked cluster constraint, such as *{ʃKPT}{ʃ}.)

I can now proceed to illustrate the constraints whose role is visible in the medial clusters of Catalan. In the first place, the full range of (phonemically contrastive) onset (C_2) consonants is found, which is to say that a positional faithfulness constraint like IDENTONSET{ʃKPT} will be highly ranked. More accurately, the C_2 position where place is preserved is not necessarily an onset, but may also be in a word-final coda, as in *temps* [tems] 'weather', *exempt* [əgzemt] 'exempt', *verb* [bɛrp] 'verb', *ponx* [pɔnʲʃ] 'punch'. That is to say, taking a licensing-by-cue approach, place is preserved where there are, in principle, adequate acoustic cues to identify it (see Chapter 5 above, and Padgett 1995b; Steriade 2001a; 2001b). In Catalan there are 'adequate cues' to place both in onsets before continuant sonorants (liquids, glides, vowels), and word-finally. As we shall see, assimilation may take place in word-final consonants before consonantal onsets, but there is no word-final place neutralization in the absence of assimilation to a following consonant. So the relevant faithfulness constraint

observed in Catalan is actually IDENTPlace-cue{ʃKPT}, abbreviated here as IDENTCue{ʃKPT}.

6.2.1 *Heterorganic nasal coda clusters*

In (2e) above are illustrated the only morpheme-internal *heterorganic* nasal coda clusters of Catalan, namely, those in which a labial nasal is followed by a denti-alveolar consonant, /mt, md, mn, ms, mz, mr/, for example, *solemne* [mn] 'solemn', *ambdós* [md] 'both'. The heterorganic nasal clusters [ŋm] (2g) and [ŋn] (2h) are less formal realizations of [gm], [gn]; I return to them below at (12), accounting here for the formal versions alone. The two cluster constraints which together rule out all but one type of heterorganic nasal+C cluster, namely, labial+denti-alveolar, are *{ʃKPT}$_{[+nas]}${ʃKP}, which bars any heterorganic nasal except before a denti-alveolar, and *{ʃK}$_{[+nas]}${ʃKPT}, which bars alveolo-palatal and velar nasals before any heterorganic consonant. The relevant hierarchy is given in (7).

(7) IDENTCue{ʃKPT} » *{ʃKPT}$_{[+nas]}${ʃKP}, *{ʃK}$_{[+nas]}${ʃKPT} » IDENT{ʃKPT}

The constraint ranking in (7) means that for all medial NC clusters markedness outranks faithfulness except in the case of /mT/, in which, before the least marked C_2 place, only the least marked heterorganic C_1 nasal place (i.e. labial) survives.

6.2.2 *Heterorganic denti-alveolar–C clusters*

Let us consider now heterorganic denti-alveolar–C clusters such as are illustrated in (2a–d) above, though I leave discussion of 'minor place' within denti-alveolar–denti-alveolar sequences (2b) for consideration later (§6.5). Heterorganic clusters of the /db/, /dg/ types are attested here, with labials or velars as C_2, but not /d̪ʒ/. That is, not even the least marked stop place, denti-alveolar, survives before the most marked alveolo-palatal C_2. The relevant constraint here is thus *{ʃKPT}$_{[-cont]}${ʃ}, outranking faithfulness (IDENT{ʃKPT}). In less formal styles, where *adverbi* 'adverb', *atmosfera* 'atmosphere', and *adquirir* 'acquire.INF' are pronounced with geminate [bb], [mm] and [kk] respectively, the active marked-cluster constraint is *{ʃKPT}$_{[-cont]}${ʃKP}. This constraint is a more general version of the constraint *{ʃKPT}$_{[+nas]}${ʃKP} mentioned in (7); it also includes the effects of *{ʃKPT}$_{[-cont]}${ʃ} just mentioned, while accommodating the clusters found in the (2c) examples, involving alveolar liquid ([+cont]) codas: *Elx* [lʃ] (toponym), *àlgebra* [lʒ] 'algebra', *xarxa* [rʃ] 'net', *conserge* [rʒ] 'concierge' (see (6) above). With this stronger constraint *{ʃKPT}$_{[-cont]}${ʃKP} in place, in less formal styles the only admissible heterorganic denti-alveolar coda consonants are continuants, namely, r/ɾ, ɫ, s/z, which is correct. In such styles where *adverbi* has [bb] rather than [db], heterorganic /bd/, /gd/, etc., survive, so faith-

fulness still outranks the most general marked-cluster constraint available in (4) *{ʃKPT}{ʃKPT}.

Tableau (8) shows some potential heterorganic inputs with a denti-alveolar in C_1 place, including nasals, to illustrate also the effect of the constraints in (7). The favoured outputs in (8) reflect a formal style with only obligatory, not optional assimilation. In (8a, b) we see heterorganic clusters surviving that are faithful to the input. In (8c–e), marked-cluster constraints rule out heterorganic outputs whatever the input.

(8)

	Input	Output	IDENTCue{ʃKPT}	*{ʃKPT}$_{[-cont]}${ʃ}	*{ʃKPT}$_{[+nas]}${ʃKP}	IDENT{ʃKPT}
(a)	db	☞ db				
		dd	*!			*
		bb				*!
(b)	dg	☞ dg				
		gg				*!
		dd	*!			*
(c)	d̪ʒ	d̪ʒ		*!		
		☞ dʲʒ				*
		d̪ð̪	*!			*
(d)	nb	nb			*!	
		☞ mb				*
		nd	*!			*
(e)	ng	ng			*!	
		☞ ŋg				*
		nd	*!			*

*{ʃKPT}$_{[-cont]}${ʃ} » IDENT{ʃKPT}

*{ʃKPT}$_{[+nas]}${ʃKP} » IDENT{ʃKPT}

6.2.3 *Labial–alveolo-palatal cluster*

I mentioned above under (2) the marked labial–alveolo-palatal cluster attested in *objecte* 'object'. If [bʒ] is taken to be morpheme-internal, such examples are evidence that IDENT{ʃKP}$_{[-nas]}$ (preserving labial stop place) outranks a marked-cluster constraint such as *{ʃKPT}$_{[-cont]}${ʃ}. IDENT{ʃKP}$_{[-nas]}$ is in turn outranked by *{ʃK}{ʃKP}, which continues to penalize heterorganic velar (or palatal) codas before non-denti-alveolars.

Admitting [bʒ] through the ranking just mentioned does entail allowing also labial–velar clusters *[bg]), though, since velars are less marked than alveolo-palatals. And by 'stringency' (6), admitting heterorganic coda labial stops involves admitting heterorganic coda labial fricatives (*[vʒ]/*[fʃ] and [vɣ]/ [*fk][7]). However, *[bg] may in turn be excluded by a constraint specific to non-continuants *{ʃKP}$_{[-cont]}${ʃK}$_{[-cont]}$, ranked above faithfulness to place IDENT{ʃKP}$_{[-nas]}$.

(9)

	Input	Output	*{ʃK}{ʃKP}	*{ʃKP}$_{[-cont]}${ʃK}$_{[-cont]}$	IDENT{ʃKP}$_{[-nas]}$	*{ʃKPT}$_{[-cont]}${ʃ}	IDENT{ʃKPT}
(a) = 8c	d̪ʒ	d̪ʒ				*!	
		☞ dʲʒ					*
(b)	bd	☞ bd					
		dd			*!		*
(c)	bʒ	☞ bʒ				*	
		d̪ʒ			*!		*
(d)	bg	bg		!*			
		☞ gg			*		*
(e)	gʒ	gʒ	*!				
		dʲʒ			*		

*{ʃK}{ʃKP}, *{ʃKP}$_{[-cont]}${ʃK}$_{[-cont]}$ » IDENT{ʃKP}$_{[-nas]}$ »
*{ʃKPT}$_{[-cont]}${ʃ} » IDENT{ʃKPT}

The 'laxer' constraint hierarchy relevant to labial–C clusters (allowing morpheme-internal [bʒ] and *[fʃ]) would be as in (9). In tableau (9) IDENT-Cue{ʃKPT} is omitted, as are the candidates that violate it. From here on its high ranking is taken as given. Example (9e) illustrates the point mentioned above, that IDENT{ʃKP}$_{[-nas]}$ is itself subordinate to a marked cluster constraint penalizing heterorganic coda velars and alveolo-palatals before non-denti-alveolars.

If [bʒ] were dealt with as morphologically complex, internal [bʒ] and the other labial-dorsal clusters would be naturally excluded together. The active constraints now would be as in (10), showing the less tolerant *{ʃKP}{ʃKP} at the

[7] [vɣ] is attested, uniquely, in *afgà* [əvˈɣa] 'Afghan'.

top of the ranking. Now IDENT{ʃKP}$_{[-\text{nas}]}$ is no longer active, and the marked cluster constraints are not ranked with respect to each other.

(10)

	Input	Output	*{ʃKP}{ʃKP}	*{ʃKPT}$_{[-\text{cont}]}${ʃ}	IDENT{ʃKPT}
(a) = 8c	d̪ʒ	d̪ʒ		*!	
		☞ dʲʒ			*
(b)	bd	☞ bd			
		dd			*!
(c)	bʒ	bʒ	*!	*	
		☞ dʲʒ			*
(d)	bg	bg	*!		
		☞ gg			*

*{ʃKP}{ʃKP}, *{ʃKPT}$_{[-\text{cont}]}${ʃ} » IDENT{ʃKPT}

6.2.4 *Alveolo-palatal–C clusters*

As far as the unattested alveolo-palatal–C clusters are concerned, the constraint hierarchy illustrated in tableau (10) also rules out those with velars or labials as C_2 (by *{ʃKP}{ʃKP}). The additional relevant cluster constraint that excludes denti-alveolar C_2 is *{ʃ}{ʃKPT} as is illustrated in tableau (11). The constraint *{ʃ}{ʃKPT} penalizes a cluster in which the most marked (alveolo-palatal) place is followed by any other place at all; it is crucial for (11b) and (11d).

(11)

	Input	Output	*{ʃ}{ʃKPT}	*{ʃKP}{ʃKP}	IDENT{ʃKPT}
(a)	ʒb	ʒβ	*!	*	
		☞ zβ			*
(b)	ʒn	ʒn	*!		
		☞ zn			*
(c)	ʒg	ʒɣ	*!	*	
		☞ zɣ			*
(d)	ɲd	ɲd	*!		
		☞ n̪d			*

*{ʃ}{ʃKPT} » IDENT{ʃKPT}

6.2.5 *Velar–C clusters*

Lastly I consider the velar–C sequences whose attested heterorganic clusters were illustrated in (2g, h), for example, *enigma* [gm] ∼ [ŋm] 'enigma' and *exacte* [əgˈzak̟tə] 'exact'. Here it proves necessary to distinguish nasal from non-nasal onsets. Velar–labial nasal [gm] ∼ [ŋm] is attested, but there is no matching cluster with a non-nasal (*gb, *kp, *kf). Proceeding on the basis of the stricter cluster constraints in (10) and (11) (thus excluding morpheme-internal [bʒ] and the other labial-dorsal clusters) I need only to introduce a more specific version of the *{ʃKP}{ʃKP} constraint such as to limit it to non-nasal C_2s, that is, *{ʃKP}{ʃ KP}$_{[-nas]}$. This is illustrated in tableau (12), which again corresponds to a formal style of pronunciation.

Strictly speaking, the modified constraint *{ʃKP}{ʃKP}$_{[-nas]}$ seen in (12) would allow *bŋ to surface; and nothing mentioned so far rules out *dŋ (or, for that matter, homorganic *gŋ). But *ŋ is not a contrastive segment in Catalan, and never appears in morpheme internal onsets, or other than before underlying velar stops /k, g/. So in any case a constraint like *ŋ is needed. It must stand below marked cluster constraints which force homorganic [ŋ] before a velar, namely, *{ʃKPT}$_{[+nas]}${ʃKP} in (8), and, of course, above faithfulness to place IDENT{ʃKPT}. That is, the relevant constraint ranking will be as (13) (but see §6.3 below for an alternative account).

(12)

	Input	Output	*{ʃ}{ʃKPT}	*{ʃKP}{ʃKP}$_{[-nas]}$	*{ʃKPT}$_{[-cont]}${ʃ}	ID{ʃKPT}
(a)	gb	gb		*!		
		☞ bb				*
(b)	gd	☞ gd				
		dd				*!
(c)	gʒ	gʒ		!*	*	
		☞ dʲʒ				*
(d)	gm	☞ gm				
		bm				*!

*{ʃKP}{ʃKP}$_{[-nas]}$ » IDENT{ʃKPT}

(13) *{ʃKPT}$_{[+nas]}${ʃKP} » *ŋ » IDENT{ʃKPT}

6.2.6 *Constraints for formal style*

Depending on whether morpheme-internal [bʒ] is to be excluded or not, alternative constraints and rankings have now been established for formal styles, and are set out here in (14) and (15).

(14) IDENTCue{ʃKPT} » *{ʃ}{ʃKPT}, *{ʃKP}{ʃKP}$_{[-nas]}$, *{ʃKPT}$_{[+nas]}${ʃKP}, *{ʃK}$_{[+nas]}${ʃKPT}, *{ʃKPT}$_{[-cont]}${ʃ} » *ŋ » IDENT{ʃKPT} » *{ʃKPT}{ʃKPT}; (*bʒ)

(15) IDENTCue{ʃKPT} » *{ʃ}{ʃKPT}, *{ʃK}{ʃKP}$_{[-nas]}$, *{ʃKP}$_{[-cont]}${ʃK}$_{[-cont]}$, *{ʃKPT}$_{[+nas]}${ʃKP}, *{ʃK}$_{[+nas]}${ʃKPT} » IDENT{ʃKP}$_{[-nas]}$ » *{ʃKPT}$_{[-cont]}${ʃ} » *ŋ » IDENT{ʃKPT} » *{ʃKPT}{ʃKPT}

The difference between the two approaches is illustrated in (16), where the grids show which heterorganic place clusters are admitted and excluded. In these diagrams, more marked clusters lie towards the top left, and less marked clusters towards the bottom right. In the more 'natural' (16a) the excluded clusters form a connected area, while in (16b) they do not.

(16)

a.

	ʃ	K	P	T
ʃ		✗	✗	✗
K	✗		✗	✓
P	✗	✗		✓
T	✗	✓	✓	

Grid corresponding to clusters admitted and barred by (14)

b.

	ʃ	K	P	T
ʃ		✗	✗	✗
K	✗		✗	✓
P	✓	✗		✓
T	✗	✓	✓	

Grid corresponding to clusters admitted and barred by (15)

While most clusters involving an alveolo-palatal C_1 are ruled out by marked heterorganic cluster constraints (*{ʃ}{ʃKPT} (11)), Catalan in fact displays no morpheme-internal preconsonantal alveolo-palatals at all, i.e. no homorganic alveolo-palatal clusters, except for the geminate palatal lateral, as in *espatlla* [əsˈpaʎʎə] 'shoulder'. Balearic and Valencian Catalan lack even this, having a non-palatal geminate instead: *espatla* Balearic [əsˈpallə]. That is, in Catalan there are no clusters *ʒɲ, *ʒʎ, *ɲʒ, *ɲɲ,[8] *ɲʎ, *ʎʒ, or *ʎɲ. Evidently, alveolo-palatal place is too marked to appear in internal codas at all. Even in affricates and 'homorganic' nasal-alveolo-palatal clusters, as in *esquitxar* [tʲ͡ːʃ] 'splash.INF', *calitja* [dʲ͡ːʒ] 'haze', *manxa* [nʲʃ] 'bellows', *diumenge* [nʲʒ] 'Sunday', the first element is palatalized alveolar, rather than strictly alveolo-palatal, according to Recasens & Pallarès (2001b). This suggests that a context-free markedness constraint against strictly alveolo-palatal place, *[−ant, +distr, −high] (abbreviated *{ʃ}), stands below IDENTCue{ʃKPT}, but above IDENT{ʃKPT}. For all practical effects context-free *{ʃ} can replace the heterorganic cluster constraint *{ʃ} {ʃKPT} in the rankings of (14) or (15). Furthermore, [ʎ] and [ɲ] do not appear in internal clusters as C_2, either. That is, there is no *lɲ, *lʎ, *zɲ, *zʎ, *rɲ, *rʎ. No doubt this reflects the inherently marked nature of alveolo-palatal sonorants in particular. Pending further investigation of this issue, I invoke

[8] 'Geminate integrity' that admits [ʎʎ] will admit [ɲɲ] also. In the present account, the lexical absence of [ɲɲ] in varieties that allow [ʎʎ] is taken to be fortuitous.

specific marked cluster constraints *Cʎ, *Cɲ, whose rank is above input–output faithfulness constraints, but is otherwise indeterminate.

6.3 NASAL AND LATERAL ASSIMILATION AND ASSOCIATED ASSIMILATIONS IN LESS FORMAL STYLE

In Catalan medial clusters, the coda contrasts of obstruent stop vs. nasal sonorant, and of obstruent stop vs. lateral sonorant, are displayed variably. In the less formal style of speech in which dental stops are assimilated in place to following labials or velars, stops of all places also become nasal before a nasal, and dental stops become lateral before a lateral. Examples given in (2a, b, g, h) are repeated here.

atmosfera [dm] ~ [mm] 'atmosphere'	*enigma* [gm] ~ [ŋm] 'enigma'
ètnic [dn̪] ~ [nn] 'ethnic'	*signar* [gn] ~ [ŋn] 'sign.INF'
hipnosi [bn] ~ [mn] 'hypnosis'	*atleta* [dɫ̪] ~ [ɫɫ] 'athlete'

In this less formal variety a more general marked cluster constraint *{ʃKPT}$_{[-\text{cont}]}${ʃKP} stands in place of the two constraints *{ʃKPT}$_{[-\text{cont}]}${ʃ}, *{ʃKPT}$_{[+\text{nas}]}${ʃKP} mentioned previously (14, 15). That is to say, in $[-\text{cont}]_1[-\text{cont}]_2$ clusters with a nasal or a lateral in C_2 position, dissimilar nasality or laterality in C_1 is more marked than similar manner of articulation maintained throughout the cluster. Note that while [dn] is more marked than [nn], [nn] is more marked than [nd]. This scale of markedness, in fact, follows from the Syllable Contact law (SYLCON; see §3.1 (8)), by which codas are preferably more sonorous than following onsets, or at least not less sonorous. The same scale dn > nn > nd also reflects positional faithfulness of onsets (by which, for input /dn/, output [nn] is more harmonic than output [dd]). *Signar* 'sign.INF' with [ŋn], by the way, shows that the ranking *ŋ » IDENT{ʃKPT} in (13) cannot be right, for in *signar* it is better to have a velar nasal before a non-velar than to lose velar place. It is strictly onset *ŋ that should be penalized, with *ŋOnset ranking above IDENT-Cue{ʃKPT}. To effect nasal and lateral assimilation it is sufficient that SYLCON outrank IDENT{ʃKPT}. In tableau (17) only marked cluster constraints relevant to nasal and lateral onsets are included. The constraint hierarchy assumed is that of (15), but the same approach applies with that of (14), *mutatis mutandis*.

(17) Nasal and lateral coda assimilation in less formal style

(a)

***atm**osfera*	SYLCON	ID{ʃKP}	*{ʃKPT}$_{[-\text{cont}]}${ʃKP}	ID{ʃKPT}	IDNas	IDLat
dm	*!		*			
bm	*!		*	*		
nm			*!		*	
☞ mm				*	*	

		SylCon	Id{ʃKP}	*{ʃKPT}[−cont]{ʃKP}	Id{ʃKPT}	IdNas	IdLat
(b)	*ètnic*						
	dn	*!					
	☞ nn					*	
(c)	*hipnosi*						
	bn	*!					
	☞ mn					*	
	nn		*!		*	*	
(d)	*enigma*						
	gm	*!		*			
	mm		*!		*	*	
	☞ ŋm			*		*	
(e)	*signar*						
	gn	*!					
	☞ ŋn					*	
	nn		*!		*	*	
(f)	*atleta*						
	dɫ	*!					
	☞ ɫɫ						*

SylCon, Ident{ʃKP} » *{ʃKPT}[−cont]{ʃKP} » Ident{ʃKPT}, IdentNas, IdentLat

All the examples in (17) show that SylCon outranks IdentNas and IdentLat. Examples (17c–e) show Ident{ʃKP} outranking the most general marked cluster constraint to allow heterorganic velar nasals in a coda. These examples require some revisions in the detail of the constraints standing above Ident {ʃKP}, as shown in table (18) below. The constraint ranking Ident{ʃKP} » *{ʃKPT}[−cont]{ʃKP} penalizes only heterorganic denti-alveolar (T) non-continuant codas, other heterorganic place being preserved by Ident{ʃKP}.

In the less formal variety, in which nasal and lateral assimilation of stops is seen, and where there are no heterorganic denti-alveolar [−cont] codas, the constraints of interest are as in table (18), with some key examples demonstrating the ranking.

6.4 INTER-WORD CODA–C CLUSTERS

I now turn to consider consonant contact in inter-word sandhi environments. Since all consonant places of articulation are available in both word-final con-

(18) Hierarchy of constraints governing heterorganic major place intra-word contacts in non-formal style in continental Catalan

Hierarchy		
IDCue{ʃKPT}	» *{ʃK}{ʃKP}$_{[-nas]}$	/gb/ [bb] ≻ *[gg]
»		
*{ʃ}{ʃKPT}	» IDENT{ʃKP}	/ɲd/ [nd] ≻ *[ɲd], /ʒb/ [zβ] ≻ *[ʒβ]
*{ʃK}{ʃKP}$_{[-nas]}$	» IDENT{ʃKP}	/gʒ/ [dʲʒ] ≻ *[gʒ], /gb/ [bb] ≻ *[gb]
*{ʃKP}$_{[+nas]}${ʃK}	» IDENT{ʃKP}	/mg/ [ŋg] ≻ *[mg], /mʒ/ [nʲʒ] ≻ *[mʒ]
*{ʃKP}$_{[-cont]}$ {ʃK}$_{[-cont]}$	» IDENT{ʃKP}	/bg/ [gg] ≻ *[bg]
»		
IDENT{ʃKP}	» SYLCON	/bʒ/ [bʒ] ≻ [d͡ʲʒ]
	» *{ʃKPT}$_{[-cont]}${ʃKP}	/bʒ/ [bʒ] ≻ [dʲʒ], /gm/ [ŋm] ≻ [mm] (17)
	» *{ʃKPT}{ʃKPT}	/gn/ [ŋn] ≻ [nn] (17), /md/ [md] ≻ [nd]
»		
SYLCON	» IDENT{ʃKPT}, IDNas	/dm/ [mm] ≻ [bm], [dm], /dn/ [nn] ≻ [dn] (17)
	» IDLat	/dl/ [ll] ≻ [dl] (17)
	» *{ʃKPT}{ʃKPT}, IDNas	/gn/ [ŋn] ≻ [gn] (17)
»		
*{ʃKPT}$_{[-cont]}${ʃKP}	» IDENT{ʃKPT}	/db/ [bb] ≻ [db], /dg/ [gg] ≻ [dg], /nb/ [mb] ≻ [nb]
»		
IDENT{ʃKPT}	» *{ʃKPT}{ʃKPT}	/zg/ [zɣ] ≻ [ɣɣ], /Rʒ/ [rʒ] ≻ [rʲʒ]
»		
*{ʃKPT}{ʃKPT}		
IDENTNasal		
IDENTLateral		

sonants and word-initial consonants in Catalan, there is considerable scope in principle for assimilation, in which markedness constraints such as have been illustrated in §6.3 would outrank correspondence. *Obligatory* place assimilation in inter-word clusters is in fact restricted to three types:

(a) minor place assimilation (e.g. bilabial to labiodental before labiodental)—which is also found within morphemes (see §6.5);
(b) coarticulation of dorsals before coronals (as also within morphemes);
(c) coarticulation/gestural blending in clusters involving denti-alveolar and alveolo-palatal place (in either order).

Assimilation of type (c) is not evidenced within morphemes because alveolo-palatals do not occur in preconsonantal codas, nor (except as discussed in §6.2) do denti-alveolars occur before alveolo-palatals. These obligatory place assimilations correspond to the cells in Table 6.1 (above) with a blank in line 1. Most of these issues will be taken up when I turn to assimilation of 'minor place' in §6.5. The general maintenance of major place contrasts in word-final codas that is

observed in continental Catalan reflects the pattern whereby output–output correspondence constraints outrank the marked cluster constraints. That is, the place contrasts displayed in phrase-final, or utterance-final, position are maintained before consonants within phrases, even when such resulting heterorganic clusters are of the type that is inadmissible within morphemes. The type of output–output correspondence constraint required is as in (19). Note that major place is referred to, namely Alveolo-palatal, Velar, Denti-alveolar, and Labial.

(19) IDENTPWd-PPhr{ʃKPT} (IDENTW-P{ʃKPT}): A segment in a phonological phrase has the same major place of articulation as its correspondent in a phonological word.

Observe the following phrases in a formal style of pronunciation (obligatory assimilation only; data after Recasens 1993: 180–84).

gat mort [dm] 'dead cat'	*set llums* [dʲʎ] 'seven lights'
cap joc [bʒ] 'no game'	*beuen llet* [nʲʎ] 'they.drink milk'
escrivint-vos [nb] 'writing-to.you'	*soldat xinès* [t͡ʲːʃ] 'Chinese soldier'

In the examples in the left-hand column there is no assimilation, even of [nb], a heterorganic cluster inadmissible within words. The right-hand column, though, shows assimilation of denti-alveolars to palatals in violation of Word–Phrase correspondence. The constraint IDENTW-P{ʃKPT} dominates the lowest-ranked marked cluster constraint of §6.2 for formal styles, namely, $*\{ʃKPT\}_{[+nas]}\{ʃKP\}$. However, the obligatory palatalization of word-final denti-alveolar stops and nasals before alveolo-palatals reveals the hierarchy in (20), where another general constraint of intra-morpheme contact in formal styles, $*\{ʃKPT\}_{[-cont]}\{ʃ\}$, continues to dominate over Word–Phrase faithfulness as it does over input faithfulness. But as within words, faith to non-denti-alveolar coda place generally outranks assimilation to alveolo-palatal place. Thus the active W-P constraint is IDENTW-P{ʃKP}.

(20) IDENTW-P{ʃKP} » $*\{ʃKPT\}_{[-cont]}\{ʃ\}$ » IDENTW-P{ʃKPT} » $*\{ʃKPT\}_{[+nas]}\{ʃKP\}$

Tableau (21) illustrates the absence or presence of place assimilation in the phrases mentioned above. In this most formal style, the only inter-word major place assimilation that is observed is of the least marked denti-alveolar /t, d, n/ to the place of a following alveolo-palatal (most marked place). As before, high-ranking IDENTCue{ʃKPT} preserving onset place (among other things) is taken as given, as are MAX, DEP, and voice assimilation. Input–output faithfulness ranks below $*\{ʃKPT\}_{[+nas]}\{ʃKP\}$.

(21) Inter-word place assimilation in the most formal style[9]

gat mort ['gat̪][ˈmɔrt̪]	IDW-P{ʃKP}	*{ʃKPT}[–cont]{ʃ}	IDW-P{ʃKPT}	*{ʃKPT}[+nas]{ʃKP}
☞ d̪m̪				
bm			*!	
cap joc ['kap]['ʒɔk]				
☞ bʒ		*		
dʲʒ	*!		*	
escrivint-vos [əskɾiˈβin][bus]				
☞ nb				*
mb			*!	
set llums ['sɛt̪][ˈʎums]				
d̪ʎ		*!		
☞ dʲʎ			*	
beuen llet ['bɛwən]['ʎet]				
nʎ		*!		*
☞ nʲʎ			*	
soldat xinès [sut̪ˈd̪at̪][ʃiˈnɛs]				
t̪ʃ		*!		
☞ tʲ͡ːʃ			*	

IDENTW-P{ʃKP} » *{ʃKPT}[–cont]{ʃ} » IDENTW-P{ʃKPT} » *{ʃKPT}[+nas]{ʃKP}

6.4.1 *Inter-word coda–C clusters in less formal styles*

Non-obligatory place and lateral assimilation of coda consonants affects consonants with the least marked, denti-alveolar place; non-obligatory assimilation of nasality affects stops of all places (Wheeler 1979: 294–306). Only slightly less than totally obligatory is the assimilation of denti-alveolar (but no other) nasal place to the place of a following consonant, as in (22), where the orthography distinguishes word-final place contrasts.

[9] Previously, we saw in version (15) of the heterorganic cluster constraint ranking the order *{ʃKPT}[+nas] {ʃKP} » IDENT{ʃKP}[–nas] » *{ʃKPT}[–cont]{ʃ}, which is inconsistent with the ranking now presented in (21). The evidence of (21) shows that, for that version of the ranking, *{ʃKPT}[+nas]{ʃKP} must be split into two separate constraints, *{ʃKPT}[+nas]{ʃKP} standing

(22) *són molts* [mm] 'they are many' *som grossos* [mg] 'we are large'
són pocs [mp] 'they are few' *any difícil* [ɲd̪] 'difficult year'
són grossos [ŋg] 'they are large'

In this less than extremely formal style, /n/ assimilates not only to alveolo-palatal place, as in (21), but also to velar and labial place; other place contrasts are maintained. That is, only the least marked nasal place assimilates to all the others. The marked cluster constraint $*\{\text{ʃKPT}\}_{[+\text{nas}]}\{\text{ʃKP}\}$ stands between IDENT W-P{ʃKP} (which preserves major place other than denti-alveolar) and IDENTW-P{ʃKPT} (which preserves all place). Thus, as in formal styles for intra-morphemic realization, $*\{\text{ʃKPT}\}_{[-\text{cont}]}\{\text{ʃ}\}$ and $*\{\text{ʃKPT}\}_{[+\text{nas}]}\{\text{ʃKP}\}$ stand together, dominating faithfulness.

A general non-formal style of utterance, in line with the regular realization of word-internal clusters, involves place assimilation of denti-alveolar stops (23a), and in addition assimilation of manner of an obstruent stop to a following nasal, or of a dental stop to lateral (23b). The orthography shows the underlying place and manner. Thus in this style, as well as place assimilation of denti-alveolars, we find assimilation of nasality in stops and of laterality too, where possible (in denti-alveolar stops, lateral labials being impossible and lateral velars rare, and perhaps never the outcome of assimilation).

(23) (a) *rat-penat* [pp] 'bat' *sud-coreà* [kk] 'South Korean'
marit feliç [pːf] 'happy husband'

(b) *en sap molt* [mm] 'knows a lot about it' *criat nadiu* [nn] 'native servant'
set nyanyos [ɲɲ] 'seven bumps' *set làmines* [ɬɬ] 'seven engravings'
cap noi [mn] 'no boy'
en trec molt [ŋm] 'I get a lot from it'

In the non-formal style, the constraints illustrated in (17), namely SYLCON, $*\{\text{ʃKPT}\}_{[-\text{cont}]}\{\text{ʃKP}\}$, also stand between IDENTW-P{ʃKP} and IDENTW-P{ʃKPT}, with IDENTW-PNas, and IDENTW-PLat lying below. That is to say, the hierarchy relevant to inter-word contacts in these contexts is the same as that for intra-word contacts, and the Word–Phrase correspondence constraints mimic input–output (I-O) faithfulness constraints (or are identical with them). In inter-word contacts a greater range of SYLCON-violating sequences is found, such as /b.r, b.l, ɲ.r/, and so on. I am assuming that appropriate faithfulness constraints, such as IDENTRhotic, IDENTW-P[cont], or segment markedness constraints (such as *DORSAL[+lat], *LABIAL[+lat]) are ranked higher than SYLCON, as is also the constraint preserving word-final consonants from incorporating into a following syllable, provided it already has an onset (i.e. ONSET » UNIFORMσ; see §3.1.1 (15), (16)).

where it is shown in (21) and $*\{\text{ʃKP}\}_{[+\text{nas}]}\{\text{ʃKP}\}$ penalizing only non-denti-alveolar heterorganic nasal place, standing above IDENT{ʃKP} » $*\{\text{ʃKPT}\}_{[-\text{cont}]}\{\text{ʃ}\}$ in (15). (In version (14) of the heterorganic cluster constraint ranking, $*\{\text{ʃKPT}\}_{[+\text{nas}]}\{\text{ʃKP}\}$ was not strictly ranked in relation to $*\{\text{ʃKPT}\}_{[-\text{cont}]}\{\text{ʃ}\}$, so IDENTW-P{ʃKPT} can simply be put between them, as it appears in (21).)

6.4.2 *Summary of style-reflecting place assimilations in inter-word consonantal contacts*

- Most formal style: only least marked denti-alveolar /t, d, n/ assimilates to the place of a following alveolo-palatal (most marked place) (21).
- Slightly less formal style: /n/ (least marked nasal place) assimilates not only to alveolo-palatal place as in (21), but also to velar and labial (22).
- General non-formal style: place assimilation of all denti-alveolar stops (23a), and in addition assimilation of manner of a denti-alveolar stop to a lateral, and of any obstruent stop to a following nasal (23b).

6.4.3 *Heterorganic major place inter-word contacts*

The remaining heterorganic major place inter-word contacts—most of the greyed-in cells of Table 6.1 above—are realized faithfully, or are subject only to minor place assimilation. These are the contacts involving a coda with any place other than the least marked denti-alveolar. These unmodified contacts, which are not attested morpheme-internally, are the following, with examples:

Labial–Alveolo-palatal
cap joc [ˈkab ˈʒɔk] 'no game', *cap lleó* [ˈkab ʎəˈo] 'no lion'
Labial–Velar
cap got [ˈkab ˈgɔt] 'no glass'
Alveolo-palatal–Labial
mateix martell [məˈteʒ mərˈteʎ] 'same hammer', *aquell foc* [əˈkɛʎ ˈfɔk] 'that fire'
Alveolo-palatal–Denti-alveolar
aquell temps [əˈkɛʎ ˈtems] 'that weather', *mateix dia* [məˈteʒ ˈðiə] 'same day'
Alveolo-palatal–Velar
mateix got [məˈteʒ ˈɣɔt] 'same glass', *company graciós* [kumˈpaɲ ɡrəˈsjos] 'amusing companion'
Velar–Labial
sac buit [ˈsag ˈbujt] 'empty bag'
Velar–Alveolo-palatal
conec lleons [kuˈnɛg ʎəˈons] 'I know lions', *dic ximpleries* [ˈdik ʃimpləˈriəs] 'I talk nonsense'

It is in these heterorganic clusters that Word–Phrase faithfulness evidently outranks the dominant marked cluster constraints that are active word-internally to exclude such clusters within morphemes. Thus, Word–Phrase faithfulness to all but the least marked denti-alveolar place, IDENTW-P{ʃKP}, outranks the

(24) Hierarchy of constraints governing heterorganic major place contacts (intra-word and inter-word) in non-formal style in continental Catalan, incorporating (18)

		Examples of inter-word heterorganic clusters preserved
IDENTCue{ʃKPT}		
»		
IDENTW-P{ʃKP}	» *{ʃ}{ʃKPT}	/ɲ#b/ [ɲb] ≻ [mb], /ʒ#d/ [ʒð] ≻ [zð]
	» *{ʃK}{ʃKP}$_{[-nas]}$	/g#b/ [gb] ≻ [bb]
	» *{ʃKP}$_{[+nas]}${ʃK}	/m#ʒ/ [mʒ] ≻ [nʲʒ]
»		
*{ʃ}{ʃKPT}		
*{ʃK}{ʃKP}$_{[-nas]}$		
*{ʃKP}$_{[+nas]}${ʃK}		
*{ʃKP}$_{[-cont]}${ʃK}$_{[-cont]}$		
»		
IDENT{ʃKP}		
»		
SYLCON		
»		
*{ʃKPT}$_{[-cont]}${ʃKP}		
»		
IDENT{ʃKPT}		
»		
*{ʃKPT}{ʃKPT}		
IDENTNasal		

heterorganic cluster constraints such as *{ʃK}{ʃKP}$_{[-nas]}$ and *{ʃ}{ʃKPT}, which, as I showed above in (9), themselves dominate IDENTI-O{ʃKP}. Thus the framework for the constraint hierarchy relevant to inter-word contacts is given in (24), representing an informal style with general place assimilation of coda denti-alveolar non-continuants, nasal assimilation of stops, and lateral assimilation of /t, d/.

6.5 MINOR PLACE ASSIMILATION

In addition to assimilation or otherwise of major place of articulation (Labial, Denti-alveolar, Alveolo-palatal, and Velar), Table 6.1 also illustrates assimilatory effects within these classes, for example, the anticipatory assimilation of bilabial to labiodental: *som feliços* [ˈsomɲ fəˈlisus] 'we are happy'. Inter-word contacts in Catalan provide many examples in particular of coronal + coronal contacts both within and between the denti-alveolar and alveolo-palatal regions. Such contacts have not been widely researched within the general phonological literature, but the Catalan facts with respect to lingual articulation have been the

subject of detailed investigation by Recasens and his colleagues. In Table 6.2 I set out some of the place distinctions made by Recasens et al., in terms of relatively familiar features. I do so not with the intention of arguing for the correctness of a particular set of features, but rather to help focus on the similarities and differences among coronals that Recasens' approach deals with. (Recasens et al. do not themselves make use of distinctive features.)

In the approach to consonant contact of Recasens (1993: 194–6; Recasens & Pallarès 2001a; 2001b: 15–16), a distinction is made between coarticulation, assimilation, and gestural blending. Coarticulation in a consonant involves the addition of an articulatory gesture without loss of the consonant's own characteristic gestures. For Recasens, the lowering of the lamino-predorsal or dorsal region of the tongue in /ʎ/ before dark [ɫ] or trilled [r], giving [lʲ], is an example of coarticulation, as is the fronting of a velar before an anterior lingual consonant, for example /gd/ → [g̟d]. Assimilation, by contrast, for Recasens does involve the loss of a characteristic gesture, though there may be addition of coarticulatory gestures also. The realization of dental /d/ as alveolar/post-alveolar [d̠] before [r], counts as assimilatory for Recasens (Recasens & Pallarès 2001b: 15). Neutralization of contrast (such as /tk/ → [kk]) necessarily involves assimilation, but not all assimilation leads to neutralization. (In the absence of neutralization, Recasens' distinction between coarticulation and assimilation is perhaps not always easy to maintain, since it depends on the notion of 'characteristic' (or *target*) gesture, which is not the same as a 'phonologically contrastive' gesture. 'Characteristic' gestures for Recasens are those realized in intervocalic consonants, more particularly, between open vowels /a_a.) Gestural blending involves gestures which are compromises between the targets of the two adjacent consonants; thus gestural blending amounts to mutual coarticulation or mutual assimilation. Recasens acknowledges that coarticulation, assimilation, and gestural blending are points on a continuum of contextual adaptation of articulation.

TABLE 6.2 Suggested feature classification for place of articulation of types represented in Table 6.1 (modified from Bonet & Lloret (1998: 127–8), in line with the observations of Recasens & Pallarès (2001b: 35–41))

	Labial		(Coronal)					(Dorsal)	
			Denti-alveolar			Alveolo-palatal		Velar	
	p b β m	f v ɱ	t d ð n̪ ɫ̪	t̠ d̠ ð̠ s z n ɾ r l ɫ		sʲ zʲ nʲ lʲ	ʃ ʒ ɲ ʎ	k̟ g̟ ɣ̟ j	k g ɣ
ant	+	+	+	+		−	−	−	−
distr	+	−	+	−		−	+	+	+
high	−	−	−	−		−	−	+	+
back	−	−	− +	−	+	−	−	−	+
	bilabial	labio-dental	lamino-dental	apico-alveolar		palatalized	alveolo-palatal	palatal	velar

In the theory of Recasens and his colleagues, the contact effects of lingual consonants depend not on articulatory place per se, but on the degree of articulatory constraint (DAC) of each of the consonants involved (Recasens et al. 1997; Recasens & Pallarès 2001a; 2001b: 16). A consonant with greater DAC has greater resistance to coarticulation, and a potentially greater effect on its neighbour, than a consonant with lower DAC. For example, [n] and [r] have similar place, but the trill requires more complex and precise articulation, i.e. [r] has a high degree of articulatory constraint (DAC). Thus in both /nr/ and /rn/ contacts the place of [r] wins, i.e. back alveolar rather than front alveolar. The DAC classification of Catalan consonants, following Recasens, is as follows:

- The grooved fricatives [s]/[z], [ʃ]/[ʒ] and the alveolar trill [r] have a **high DAC with respect both to the tongue dorsum ([−anterior]) and to the tongue front ([+anterior])**.
- Other consonants with raising of the dorsum, namely, alveolo-palatal sonorants ([ɲ], [ʎ]), palatals ([j]), velars ([k]/[g]/[ɣ]), and dark [ɫ], have a **high DAC with respect to the dorsum [−anter]**, but a low DAC with respect to the front of the tongue [+anter], either because the front of the tongue is flexible and adaptable (apical [ɫ]) or because the front of the tongue is not involved at all in the target articulation (the non-denti-alveolars).
- The other lingual consonants ([t]/[d]/[ð], [ɾ], [n]) have low DAC with respect to both the dorsum and the front of the tongue.

In addition to the effect of the DAC on contextual adaptation, Recasens & Pallarès (2001b: 17) recognize the universal predominance of regressive assimilation, which, as mentioned in §6.1, reflects the generally greater perceptibility of CV transitions (in onsets) than VC transitions (in codas). Thus, when adjacent consonants are both of high DAC, the place and/or manner of the second (onset) usually predominates. So, a cluster of two consonants of high $DAC_{[+anter]}$, such as /sr/ (or /zr/), is realized to the advantage of the C_2, [rr] ∼ [r]: *els roures* [əɫˈrowɾəs] 'the oaks'. With two consonants of high $DAC_{[-anter]}$ such as /ɲ/, /ɫ/, the first consonant displays coarticulation with the second, more than the other way round: /ɲɫ/ *any luctuós* [nʲɫ] 'tragic year', /ɫɲ/ *el nyanyo* [lɲ] 'the lump'.

For Recasens the articulatory adaptations in lingual C_1C_2 contacts are related to degrees of articulatory constraint in the following way. We find:

- **gestural blending** when neither C_1 nor C_2 has strong articulatory constraints with respect to the other;
- **assimilation** of C_1 when only C_2 has strong articulatory constraints;
- **no change, or coarticulation or progressive assimilation** when only C_1 has strong articulatory constraints (Recasens & Pallarès 2001b: 61).

This scene of competing articulatory and prosodic demands lends itself naturally to an OT approach. The aim is to identify, as far as possible, the relative strength of the constraints that give rise to particular outcomes in consonantal contacts of specific types. The range of contacts that are the focus of this section comprises

those enclosed in the bold box in Table 6.1, namely, contacts between coronal consonants (denti-alveolar or alveolo-palatal). Some of these, in fact, involve 'major place' assimilation of denti-alveolar to alveolo-palatal place or conversely. The representations in Table 6.1 do not explicitly illustrate gestural blending, so I begin by commenting on these.

A cluster of an alveolo-palatal sonorant and a denti-alveolar stop (ɲd, ɲn, ʎd, ʎn) is produced with gestural blending; that is to say, the tongue makes contact throughout over the area appropriate to both consonants (Recasens & Pallarès 2001b). So [ɲd] is actually realized as a homorganic dental-alveolo-palatal cluster, and [ɲn] as a homorganic alveolar-alveolo-palatal cluster. From a constraint perspective, I suggest that such gestural blending outputs violate neither marked (heterorganic) cluster constraints nor faithfulness to input minor place. Similarly, what are represented in the table as [n̪d] and [ɫ̪d] are gesturally blended dental-alveolar clusters (which are so realized within morphemes also); likewise [dn̪] and [dɫ̪], according to Recasens (1993: 188). In what follows, gesturally blended clusters are indicated with ^{gb}.

Sibilant+rhotic clusters /ʒr/ and /zr/ may involve some blending of *manner* of articulation, whereby the rhotic becomes an assibilated approximant (here symbolized [ɹ^{s}])—no doubt, as Recasens & Pallarès say, because of the competing requirements of precise articulation for a sibilant fricative and for a trill. In fact, the inter-word sequence /zr/, quite frequent in discourse due not least to plural /+z/, is more usually realized [r] with deletion of /z/, as mentioned above, as in *els roures* [əɫˈrowɾəs] 'the oaks'. Note that it is the production of the sequence of sibilant and rhotic in this order that appears to be articulatorily problematic; /rz, rs, rʃ, rʒ/ are realized faithful to the input. It seems that trills are best prepared from vowels (or silence), while sibilants can be prepared from more or less anywhere. Lastly, the realization of /ʒz/ as [z^{j}z^{j}] or [z^{j}] may, perhaps, be interpreted as gestural blending. (This is the only contact in Catalan with such an outcome involving superficial unfaithfulness to onset place—/z/ moves from the denti-alveolar to the alveolo-palatal category.)

I adopt for this analysis several aspects of Recasens & Pallarès's DAC theory. Specifically, s/z, r, and ʃ/ʒ are classified as DAC$_{[+ant]}$, having a high degree of articulatory constraint with respect to the tongue front ([+anterior]), while s/z, r, ɫ, ʃ/ʒ, ɲ, ʎ, and k/g/ɣ are classified as DAC$_{[-ant]}$, having a high degree of articulatory constraint with respect to the tongue dorsum ([−anterior]). The two groups together are classified as 'DAC'. The DAC classification is summarized in Table 6.3.

TABLE 6.3 DAC classification of Catalan consonants

DAC		¬DAC
DAC$_{[+ant]}$	DAC$_{[-ant]}$	
s/z, r, ʃ/ʒ	s/z, r, ʃ/ʒ ɫ, ɲ, ʎ, k/g/ɣ	p/b/β, m, f/v, t/d/ð, n

In fact, I do not pursue here the DAC effects of, or on, the velars k/g/ɣ whose coarticulatory outcomes—fronting especially before obstruents—are illustrated in the bottom row of Table 6.1.

6.5.1 *Survey of cluster realizations*

Of the sixty-four cells within the bold box in Table 6.1, corresponding to tongue front contacts, twenty-six contain clusters which are essentially homorganic, namely the nine alveolo-palatal–alveolo-palatal contacts, the sixteen alveolar–alveolar contacts, and the one dental–dental contact. Of these, only /zr/, usually realized [r], calls for any comment (see above). The others are realized faithfully. (There are eight cells in the boxed area of the table that correspond to C_iC_i sequences underlyingly. They are resolved as geminates, except that the anti-gemination markedness constraint *GEMINATESibilant (*GEMSib) prefers single sibilants, and an equivalent *GEMINATERhotic (*GEMRhot) prefers [r] to [rr].) Clusters involving a coda rhotic are not subject to any place (or manner) assimilation. Despite the articulatory constraints for trills[10] that Recasens & Pallarès discuss, coda trills neither require nor are subject to any modification of place or manner. This is perhaps not surprising, inasmuch as rhotics, high on the sonority scale, make generally good codas.

Of the remaining thirty-eight coronal clusters involving different input place,[11] the majority (nineteen) are resolved by regressive assimilation; ten are resolved by gestural blending or mutual assimilation (as also may be /zr/), two by progressive assimilation, and seven remain faithfully heterorganic. (Optional manner assimilations of nasality and laterality are ignored here.) These types are exemplified in (25–8). Input forms are given with voicing agreement already noted.

(25) Regressive assimilation
  (a) denti-alveolar–alveolo-palatal
    soldat xinès /t#ʃ/ [t͡ʲːʃ] 'Chinese soldier'
    set nyanyos /d#ɲ/ [dʲɲ] 'seven bumps'
    set llums /d#ʎ/ [dʲʎ] 'seven lights'
    són joves /n#ʒ/ [nʲʒ] 'they are young'
    un nyanyo /n#ɲ/ [ɲɲ] 'a bump'
    beuen llet /n#ʎ/ [nʲʎ] 'they drink milk'
    el joc /ɫ#ʒ/ [lʒ] 'the game'
    el nyanyo /ɫ#ɲ/ [lɲ] 'the bump'
    animal lluent /ɫ#ʎ/ [lʎ] 'glistening animal'
    dos generals /z#ʒ/ [ʒʒ] ~ [ʒ] 'two generals'

[10] Within the coronal contact sphere, all the rhotic contacts except one involve a trill in continental eastern Catalan; the tap [ɾ] is found before the voiced non-strident dental [ð].

[11] Including /zr/ for completeness, though in the classification used here /z/ and /r/ have the same place.

Among this group, observe that /n#ɲ/ and /z#ʒ/ result in geminates, not *[nʲɲ], *[zʲʒ]. Note also that the assimilated version of dark /ɫ/ is not a palatalized lateral, but a develarized one (Recasens 1993: 178–9). A develarized lateral accommodates place to an adequate degree while remaining faithful to major place (i.e. is not shifted from the denti-alveolar to alveolo-palatal category).

(b) alveolo-palatal–alveolar DAC
mateix límit /ʒ#ɫ/ [zʲɫ] 'same limit'
any luctuós /ɲ#ɫ/ [nʲɫ] 'tragic year'
company sinistre /ɲ#s/ [nʲs] 'sinister companion'
company ric /ɲ#r/ [nʲr] 'rich companion'
ball lúdic /ʎ#ɫ/ [lʲɫ] 'playful dance'
vell zoòleg /ʎ#z/ [lʲz] 'old zoologist'
tall rodó /ʎ#r/ [lʲr] 'round slice'

(c) dental–alveolar DAC
ha afegit zeros /d̪#z/ [d̪͡z] 'has added zeros'
soldat rus /d̪#r/ [d̪r] 'Russian soldier'

(26) Gestural blending (or mutual assimilation)

(a) alveolo-palatal–denti-alveolar
un any després /ɲ#d̪/ [ɲdᵍᵇ] 'a year after'
any nou /ɲ#n/ [ɲnᵍᵇ] 'new year'
gall dindi /ʎ#d̪/ [ʎdᵍᵇ] 'turkey'
de bell nou /ʎ#n/ [ʎnᵍᵇ] 'anew'

(b) dental stop–alveolar sonorant stop or vice versa
són durs /n#d̪/ [ndᵍᵇ] 'they are hard'
el test /ɫ#t̪/ [ɫtᵍᵇ] 'the flower-pot'
set nois /d̪#n/ [dnᵍᵇ] (∼ [nn]) 'seven boys'
fet lògic /d̪#ɫ/ [dɫᵍᵇ] (∼ [ɫɫ]) 'logical fact'
Note that /ɫ/ is interpreted as a stop [−cont] when adjacent to t/d (see §10.1).

(c) alveolo-palatal fricative−alveolar fricative or rhotic trill
mateix so /ʃ#s/ [sʲsʲ] ∼ [sʲ] 'same sound' (interpreted as mutual assimilation rather than gestural blending, inasmuch as the output is all [–anterior, −distr])
peix recuit /ʒ#r/ [zʲɹˢ] ([ʒrᵍᵇ]) 'recooked fish' (gestural blending of manner, producing a sibilant rhotic approximant)

(d) alveolar fricative−trill
dos reis /z#r/ [ɹˢɹˢ] ([zrᵍᵇ]) 'two kings'. This pronunciation, which is a very formal one, involves gestural blending of manner to produce a sibilant rhotic approximant, similar to the example in (26c) above. The more general pronunciation of this sequence involves deletion [ˈdoˈrejs] (or fusion [r_{12}]).[12]

[12] I am grateful to Daniel Recasens (p.c.) for discussion of the realization of these sibilant + rhotic clusters.

(27) Progressive assimilation
  (a) coronal fricative–dental
    es diu /z#d̪/ [zð̪] 'is called'
    creix tort /ʃ#t̪/ [ʃt̪] 'grows twisted'

(28) Faithful heterorganic
  (a) alveolo-palatal fricative–alveolar nasal
    mateix nombre /ʒ#n/ [ʒn] 'same number'
  (b) alveolar fricative–alveolo-palatal sonorant
    les nyores /z#ɲ/ [zɲ] 'the chillies'
    les llaunes /z#ʎ/ [zʎ] 'the cans'
  (c) alveolar rhotic–dental
    per dir /R#d/ [rð] 'in order to say'
  (d) alveolar rhotic–alveolo-palatal
    per gener /R#ʒ/ [rʒ] 'for January'
    per nyores /R#ɲ/ [rɲ] 'for chillies'
    per llaunes /R#ʎ/ [rʎ] 'for cans'

In the following exposition of the OT constraints and their ranking that the above pattern of contact resolution reveals, I start from the lowest-ranked constraints that are relevant at all, noting that in Catalan MAX and DEP are not available to resolve place contacts by deletion or epenthesis.

6.5.2 *Denti-alveolar contacts*

The examples in (26b) show that heterorganicity *within* denti-alveolar place is not tolerated. For this I propose a heterorganicity or marked cluster constraint *DISTANCEFront (*DISTFront) (29). *DISTFront is a gradient constraint that prefers among coronal clusters candidates with less distance of place between the members to clusters with greater distance.[13] One interpretation is to number the CORONAL places between 1 and 4, as follows: dental = 1, alveolar = 2, palatalized alveolar = 3, (strict) alveolo-palatal = 4. On this basis, identical CORONAL place scores 0 (difference between C_1 and C_2), and maximally displaced strict alveolo-palatal–dental scores 3 (4–1). *DISTFront outranks IDENTDental and IDENTAlveolar. Tableau (30) compares candidates for /d̪n/ (e.g. *set nois* 'seven boys') and /ɫd̪/ (e.g. *el dinar* 'the lunch'). Gesturally blended consonants are understood to be faithful to the places of both their constituents. The other gesturally blended clusters (26a) are similarly selected by *DISTFront in preference to heterorganic candidates.

[13] If gradient constraints are to be avoided as McCarthy (2003b) recommends, *DISTFront can easily be replaced by a series of constraints expressing the markedness of each place pair, inherently ranked on a scale of distance.

(29) *DISTANCEFront (*DISTFront): Score one violation mark for each degree of difference between adjacent Coronal consonants less sonorous than [ɾ][14] on the scale dental = 1, alveolar = 2, palatalized alveolar = 3, (strict) alveolo-palatal = 4. (DISTANCE violations are represented numerically rather than with asterisks, but this is merely a matter of simplifying the visual presentation.)

(30)

/d̪n̪/	*DISTFront	IDENTDental	IDENTAlveolar
d̪n̪	1!		
d̪n̪	0		*!
d̪n̪	0	*!	
☞ dn^{gb}	0		
/ɬd̪/			
ɬd̪	1!		
ɬd̪	0		*!
ɬd̪	0	*!	
☞ ɬd^{gb}	0		

*DISTFront, IDENTDental, IDENTAlveolar

By contrast to the situation in (30), when one of the consonants involved in dental–alveolar contact is a fricative (s/z) or onset /r/, the contact is resolved in favour of the alveolar (examples (25c), (27a)). Within the DAC theory of Recasens & Pallarès (2001a; 2001b), this difference follows from the fact that s/z and r (in contrast to t/d, n, ɬ) have a high degree of articulatory constraint (DAC) with respect to adjacent [+ant] consonants. In OT terms we could say that faithfulness to dental or alveolar place ([±distributed]) in DAC$_{[+ant]}$ consonants (i.e. IDENT[distr]DAC$_{[+ant]}$ (31)) outranks *DISTFront. IDENTDAC constraints are meant to preserve against gestural blending, unlike general IDENTFeature constraints.

(31) IDENT[distr]DAC$_{[+ant]}$: A consonant of the class DAC$_{[+ant]}$ (viz. s/z, ʃ/ʒ, r) has the same value for [distributed] in the output as in the input.

In tableau (32) candidates corresponding to /dz/, /dr/, /zd/, and /ʃt/ are compared: in (32a) *ha afegit zeros* /d̪#z/ [d̪:z] 'has added zeros', (32b) *soldat rus* /d̪#r/ [d̪r]

[14] [ɾ], like the other consonants of highest sonority, the glides [j] and [w], does not undergo assimilation in place or manner to an adjacent consonant or provoke assimilation of place or manner in an adjacent consonant. Generally, the place of [ɾ] is preserved by IDENTW-P{ʃKPT}(place faithfulness), but if [ɾ] were not excluded from the effects of *DISTANCEFront, the strictly heterorganic cluster [ɾð] (alveolar-dental) would be wrongly penalized and assimilated *[ɾð̪] would win. Example (32c) shows progressive assimilation to alveolar place when C_1 is [z].

'Russian soldier', (32c) *es diu* /z#d̪// [zð̪] 'is called', (32d) *creix tort* /ʃ#t̪/ [ʃt̪] 'grows twisted'.

(32)

	/d̪z/	IDENT[distr]DAC[+ant]	*DISTFront	IDENTDental	IDENTAlveolar
(a)	d̪z		1!		
	d̪z	*!	0		*
	☞ d̪z		0	*	
(b)	/d̪r/				
	d̪r		1!		
	d̪r	*!	0		*
	☞ d̪r		0	*	
(c)	/zd̪/				
	zð̪		1!		
	zð̪	*!	0		*
	☞ zð̪		0	*	
(d)	/ʃt̪/				
	ʃt̪		3!		
	☞ ʃt̪		2	*	
	sʲt̪	*!	2		

IDENT[distr]DAC[+ant] » *DISTFront » IDENTDental, IDENTAlveolar

At the end of §6.2 I mentioned the context-free markedness constraint *{ʃ}, disfavouring all [−ant, +distr, −high] consonants, that stands below IDENT-Cue{ʃKPT} (allowing preservation of alveolo-palatals in onsets), but above IDENT{ʃKPT}. The effect of *{ʃ} can be seen in the examples of (25a), where, in assimilating denti-alveolar to alveolo-palatal place, a somewhat heterorganic palatalized alveolar is preferred to a fully homorganic alveolo-palatal, as in *diumenge* [nʲʒ] 'Sunday' (better than *[ɲʒ]). Example (32d) shows that *{ʃ} must be ranked below IDENT[distr]DAC$_{[+ant]}$, otherwise [sʲt̪] would win over [ʃt̪]. But [nʲʒ] ≻ [ɲʒ] in *diumenge* shows that *{ʃ} ranks above *DISTFront which favours maximum assimilation. Underlying alveolo-palatals are acceptable in codas, *ceteris paribus*, but an alveolo-palatal is not an acceptable realization of a denti-alveolar, again *ceteris paribus*.

At (25a) I remarked that de-velarization was the acceptable strategy for assimilation of a back /ɬ/ to a following alveolo-palatal. One might ask, then, why de-velarization of /ɬ/ is not more generally an acceptable strategy in order to decrease the distance in a CORONAL cluster. It must be the case that a faithfulness constraint, say IDENTONS[back]Lat (or just IDENTONS[back]), outranks IDENT

[distr]DAC$_{[+ant]}$, so that for *mateix límit* /ʒ#ɬ/ 'same limit' [zʲɬ] is preferred to [ʒl] with the same degree of accommodation (assuming [+back] → [−back] counts as one degree less distant), or to [zʲl] with yet more accommodation. Still, [zʲ] here, and many other cases in (25b) with alveolo-palatals realized as palatalized alveolars, shows that violations of IDENT[distr]DAC$_{[+ant]}$ are tolerated.

I argue that this unfaithfulness to strict alveolo-palatal place, which in some cases (where IDENT[distr]DAC$_{[+ant]}$ is not in play) reflects the *{ʃ} constraint, is more often due to a specific version of *DISTFront that penalizes particularly distance between consonants with high DAC, which I call *DISTDAC (33). The 'articulatory constraint' that DAC refers to makes difference of place between adjacent DAC consonants more 'costly', or less acceptable than difference of place between coronal consonants in general.

(33) *DISTANCEDAC (*DISTDAC): Score one violation mark for each degree of difference between adjacent DAC Coronal consonants on the scale dental = 1, alveolar = 2, palatalized alveolar = 3, (strict) alveolo-palatal = 4.

Lastly in this group of constraints I reintroduce the Word–Phrase correspondence constraint IDENTPWd-PPhr{ʃKPT}, repeated here from (20).

(34) IDENTPWd-PPhr{ʃKPT} (IDENTW-P{ʃKPT}): A segment in a phonological phrase has the same major place of articulation as its correspondent in a prosodic word.

IDENTW-P{ʃKPT} does not penalize place accommodation within the *major* categories alveolo-palatal, velar, labial, denti-alveolar, but it does penalize accommodations that cross category boundaries (and are thus, potentially, at least, neutralizing).[15] In (35) are presented representative clusters of an alveolo-palatal fricative followed by denti-alveolar stop or non-rhotic sonorant. Example (35a) corresponds to the case of *creix tort* /ʃ#t̪/ [ʃt̪] 'grows twisted' (27a), to which consideration has already been given in (32d); (35b) corresponds to *mateix nombre* /ʒ#n/ [ʒn] 'same number' from (28); and (35c) corresponds to *mateix límit* /ʒ#ɬ/ [zʲɬ] 'same limit' from (25b).

(35)

(a)

/ʃt̪/	IDW-P{ʃKPT}	IDOns[back]	*DISTDAC	ID[distr]DAC$_{[+ant]}$	*{ʃ}	*DISTFront
i. ʃt̪					*	3!
ii. ☞ ʃt̪					*	2
iii. ʃtʲ	*!				*	1
iv. sʲt̪				*!		2

[15] IDENTW-P{ʃKPT} is probably not differently ranked at this point from IDENTI-O{ʃKPT}.

(b) /ʒn/	IDW-P{ʃKPT}	IDOns[back]	*DISTDAC	ID[distr]DAC[+ant]	*{ʃ}	*DISTFront
i. ☞ ʒn					*	2
ii. zʲn				*!		1
iii. ʒnʲ	*!				*	1
(c) /ʒɫ/						
i. ʒɫ			2+1!		*	2+1
ii. ☞ zʲɫ			1+1	*		1+1
iii. zʲl		*!	1			
iv. ʒl		*!	2		*	2

IDENTW-P{ʃKPT}, **IDOns[back] » *DISTDAC » IDENT[distr]DAC[+ant] » *{ʃ}, *DISTFront**[16]

In (35a), *DISTDAC is not relevant since both consonants are not DAC. Candidate (iii)'s onset consonant shifts category from denti-alveolar to alveolo-palatal, and so it falls to IDENTW-P{ʃKPT}, even though the cluster elements are closer, by *DISTFront, than in the alternatives. Of the two candidates, (ii) and (iv), scoring 2 on *DISTFront, the latter incurs a fatal violation of IDENT[distr]DAC[+ant], since input /ʃ/ is realized [sʲ]. In (35b), candidate (i) wins for similar reasons as candidate (ii) in (35a); in this case, however, as /n/ is already underlyingly alveolar, no better candidate than the faithful one is available. In (35c) both members of the cluster are DAC, so *DISTDAC plays a role—in the scoring of *DISTDAC violations, the difference in [±back] is scored separately, as +1, for expository reasons, but it is the total that counts. By *DISTDAC the faithful candidate (i) is eliminated. However, candidate (iii), the winner on *DISTDAC, and candidate (iv), which ties with (ii) on *DISTDAC, lose to candidate (ii) because of fatal IDENTOns[back] violation. Faithful dark [ɫ] is always preferred in an onset.

Tableau (36) considers cases of an alveolo-palatal sonorant (ɲ or ʎ) followed by a high DAC alveolar (s/z, ɫ, r), exemplified in (25b), for example, *company sinistre* /ɲ#s/ [nʲs] 'sinister companion' (36a), *company ric* /ɲ#r/ [nʲr] 'rich companion' (36b), *ball lúdic* /ʎ#ɫ/ [lʲɫ] 'playful dance' (36c).

[16] Bold type in the constraint hierarchy here and subsequently draws attention to those elements of the ranking that are demonstrated by the tableau above.

(36)

	/ɲs/	IDW-P{ʃKPT}	IDOns[back]	*DISTDAC	ID[distr]DAC[+ant]	*{ʃ}	*DISTFront
(a)	i. ɲs			2!		*	2
	ii. ☞ nʲs			1			1
	iii. ns	*!		0			0
	iv. nsʲ	*!		1		*	1
(b)	/ɲr/						
	i. ɲr			2!		*	2
	ii. ☞ nʲr			1			1
(c)	/ʎɫ/						
	i. ʎɫ			2+1!		*	2+1
	ii. ☞ lʲɫ			1+1			1+1
	iii. lʲl		*!	1			1
	iv. ʎl		*!	2		*	2

IDENTW-P{ʃKPT}, IDOns[back] » ***DISTDAC** »
(IDENT[distr]DAC[+ant] » *{ʃ},) ***DISTFront**

In (36a) again the candidates (iii, iv) with transfer across place category (neutralization of contrast, violating IDENTW-P{ʃKPT}) are eliminated, and the winner is the more assimilated (ii) of the remaining candidates. In (36c) only candidates respecting major place are considered; (iii) and (iv) fail, having develarized onsets, and the winner is again the more assimilated of the remaining candidates.

6.5.3 *Alveolo–palatal onsets*

I consider next some example clusters from (25a) with alveolo-palatal onsets. In most cases a coda denti-alveolar moves into the alveolo-palatal category to make the cluster less marked. This aspect has already been considered above, involving the constraint ranking *{ʃKPT}[−cont]{ʃKP} » IDENTW-P{ʃKPT} in (24). A general marked cluster constraint thus outranks faithfulness here. The clusters in (37) correspond to those in the examples *soldat xinès* /t#ʃ/ [t͡ʲːʃ] 'Chinese soldier' (37a), *són joves* /n#ʒ/ [nʲʒ] 'they are young' (37b), and *el joc* /ʎ#ʒ/ [lʒ] 'the game' (37c).

(37)

	/tʃ/	*{ʃKPT}[−cont]{ʃKP}	IDW-P{ʃKPT}	IDOns[back]	*DISTDAC	ID[distr]DAC[+ant]	*{ʃ}	*DISTFr
(a)	i. t̪ʃ	*!					*	3
	ii. t̪ʃ	*!					*	2
	iii. ☞ tʲʃ		*					1
	iv. cʃ		*				**!	0
	v. tʲsʲ		*			*!		0
(b)	/nʒ/							
	i. nʒ	*!					*	2
	ii. ☞ nʲʒ		*				*	1
	iii. nʲzʲ		*			*!		0
	iv. ɲʒ		*				**!	0
(c)	/ɫʒ/							
	i. ɫʒ				2+1!		*	2+1
	ii. ☞ lʒ				2		*	2
	iii. lʲʒ		*!		1		*	1
	iv. lʲzʲ		*!		0	*		0

***{ʃKPT}[−cont]{ʃKP} » IDENTW-P{ʃKPT}**
(, IDENTOnset[back]) » ***DISTDAC** »
IDENT[distr]DAC[+ant] » *{ʃ} » *DISTFront

Tableau (37) does not actually show crucial ranking of *{ʃKPT}$_{[-\text{cont}]}${ʃKP} over IDENTW-P{ʃKPT}, though evidence for this was presented above in (21). What (37b) shows, however is that *{ʃ} ranks above *DISTFront: candidate (ii) (partial assimilation) is preferable to candidate (iv) (full place assimilation). The form of the constraint, as a markedness constraint, perhaps does not make clear that its interesting effect is to prevent merger of /nʒ/ with /ɲʒ/, a fact which may suggest that faithfulness, or contrast preservation, is really at issue. Still, it is not wrong for preservation of contrast to emerge through ranking of various constraints. To put it another way, presence of the markedness constraint *{ʃ} is contributing to saying that true alveolo-palatals are only found through faithfulness to input specification, with a justifiable qualification about geminates, to be mentioned shortly.

In (39) and (40) I compare some clusters with alveolo-palatal sonorants in C_2 position—nasals in (39), laterals in (40). These have been illustrated in (25a) and (28). The types illustrated here correspond to *set nyanyos* /d#ɲ/ [dʲɲ] 'seven bumps' (39a), *un nyanyo* /n#ɲ/ [ɲɲ] 'a bump' (39b), *animal lluent* /ɬ#ʎ/ [lʎ] 'glistening animal' (40a), *beuen llet* /n#ʎ/ [nʲʎ] 'they drink milk' (40b), and *les llaunes* /z#ʎ/ [zʎ] 'the cans' (40c). Here the last of the DAC-specific constraints (38) is introduced, IDENTOnset[distr]DAC, the constraint penalizing, in an onset, deviation from input strict alveolo-palatal place, and strict alveolar place for s/z, r, ɬ.

(38) IDENTOnset[distr]DAC (IDOn[distr]DAC): A consonant of the class DAC in onset position has the same value for [distributed] in the output as in the input.

(39)

	/dɲ/	*{ʃKPT}[−cont]{ʃKP}	IDOn[distr]DAC	IDW-P{ʃKPT}	IDOn[bk]	*DISTDAC	ID[distr]DAC[+ant]	*{ʃ}	*DISTFr
(a)	i. d̪ɲ	*!						*	3
	ii. ☞ dʲɲ			*				*	1
	iii. dʲnʲ		*!	*					0
(b)	/nɲ/								
	i. nɲ	*!						*	2
	ii. nʲɲ			*				*	1!
	iii. nʲnʲ		*!	*					0
	iv. ☞ ɲɲ			*				*[17]	0

***{ʃKPT}[−cont]{ʃKP}**, IDENTOnset[distr]DAC »
IDENTW-P{ʃKPT}, IDENTOnset[back] » *DISTDAC »
IDENT[distr]DAC[+ant] » *{ʃ} » *DISTFront

In (39a) and (39b) the faithful candidates (i) are eliminated as they violate the higher-ranked marked cluster constraint *{ʃKPT}$_{[-\text{cont}]}${ʃKP} introduced in (20–24) that penalizes clusters with stops of the least marked place, denti-alveolar, preceding consonants of the most marked place, alveolo-palatal. All alternative candidates with non-faithful codas, of course, violate the correspondence constraint IDENTW-P{ʃKPT}. In each case candidate (iii) with a palatalized alveolar in onset position falls to the IDENTOnset[distr]DAC constraint introduced in (38).

[17] I am assuming a geminate incurs only one marked place violation, unlike clusters of same place but different manner.

In (39b) the geminate [ɲɲ] wins on *DistFront. As Table 6.1 indicates, a geminate alveolo-palatal nasal is an optional alternative in type (39a) also (*set nyanyos* /d#ɲ/ [dʲɲ] ∼ [ɲɲ]). Thus, when optional nasal assimilation applies by active SylCon, *DistFront prefers [ɲɲ] to [nʲɲ].

(40)

	/ɬʎ/	*{ʃKPT}[−cont]{ʃKP}	IDOn[distr]DAC	IDW-P{ʃKPT}	IDOn[bk]	*DISTDAC	ID[distr]DAC[+ant]	*{ʃ}	*DISTFr
(a)	i. ɬʎ					2+1!		*	2+1
	ii. ☞ lʎ					2		*	2
	iii. ɬʲʎ			*!		1+1		*	1+1
	iv. lʲʎ			*!		1		*	1
	v. lʲlʲ		*!	*		0			0
	vi. ʎʎ			*!		0		*	0
(b)	/nʎ/								
	i. nʎ	*!						*	2
	ii. ☞ nʲʎ			*				*	1
	iii. nʲlʲ		*!	*					0
	iv. ɲʎ			*				**!	0
(c)	/zʎ/								
	i. ☞ zʎ					2		*	2
	ii. zʲʎ			*!		1	*	*	1
	iii. ʒʎ			*!		0	*	**	0
	iv. zʲlʲ		*!	*		0	*		0

*{ʃKPT}[−cont]{ʃKP}, IDENTOnset[distr]DAC »
IDENTW-P{ʃKPT}, IDENTOnset[back] » ***DISTDAC** »
IDENT[distr]DAC[+ant] » ***{ʃ}** » ***DISTFront**

Tableau (40) shows alveolo-palatal laterals in C_2 position.

In (40) the differences between the three types (a–c) are of particular interest, given that C_2 is the alveolo-palatal lateral [ʎ] in each case, and that C_1 is an alveolar consonant. Of the three C_1s compared, only /n/ is [−cont], so (40b) is the only one of the three types subject to *{ʃKPT}[−cont]{ʃKP}, which eliminates candidates with C_1 denti-alveolar place. Thus in (40a) and (40c) candidates faithful to the major place of the input C_1 survive IDENTW-P{ʃKPT}, to which

more assimilated candidates fall. In (40c), the faithful candidate (i) is the only survivor (and [zlʲ] would be eliminated by IDENTOnset[distr]DAC). In (40a) there remains an alternative, candidate (ii), with a develarized alveolar lateral which beats (i) on the *DISTDAC constraint. Some assimilation, within the limits of faithfulness to major place, is the optimal result.

6.5.4 *High DAC contacts*

Lastly, I turn to cases involving clusters both of whose members are of the class with high $DAC_{[+ant]}$, namely, s/z, ʃ/ʒ, r. As I mentioned before, coda [r] neither undergoes nor provokes assimilations; as far as manner is concerned, I assume a high-ranked relevant IDENTTrill constraint (probably IDENTCodaTrill » IDENTTrill, or IDENTCueTrill » IDENTTrill, where trills are better cued post-vocalically or in absolute initial position than elsewhere), along with IDENTW-P{ʃKPT}. Palatalized [rʲ] is excluded by undominated *R[−ant]. So the cases under consideration can be illustrated by the examples *dos generals* /z#ʒ/ [ʒʒ] ∼ [ʒ] 'two generals' (25a), *mateix so* /ʃ#s/ [sʲsʲ] ∼ [sʲ] 'same sound', and *peix recuit* /ʒ#r/ [zʲɹˢ]) 're-cooked fish' (26c), and *dos reis* /z#r/ [ɹˢɹˢ] 'two kings' (27b). The constraint hierarchy I have presented so far does not account for the actual realization of these clusters. Rather, they predict *[zʒ], *[sʲs], *[zʲr] and *[zr] respectively. From the perspective of Recasens & Pallarès (2001b) these clusters provoke special articulatory difficulty because a consonant with a high degree of articulatory constraint with respect to the front of the tongue ([+anterior]) is adjacent to another such, and at least one of them is itself [+anterior]. I infer a constraint against 'faithful' gestural blending applies: *gb$DAC_{[+ant]}$, dominating *DIST$DAC_{[+ant]}$. The conflict is resolved in such a way that there is minimum place difference (though alveolo-palatal place must not be obliterated) and no disagreement in sibilance. Observe that the mutual assimilation/fusion of the type in *mateix so* /ʃ#s/ [sʲsʲ] ∼ [sʲ] 'same sound' is the only example in the whole of Table 6.1 where onset place is compromised, provided we interpret gestural blending as 'faithful'. Thus it is clear that the constraint favouring it outranks the highest constraint considered in this section requiring faithful denti-alveolar onsets, namely IDENTW-P{ʃKPT}. I invoke two separate heterorganic cluster constraints that focus on $DAC_{[+ant]}$ consonants specifically.

(41) *DIST$DAC_{[+ant]}$: Score one violation mark for each degree of difference between adjacent $DAC_{[+ant]}$ consonants on the scale dental = 1, alveolar = 2, palatalized alveolar = 3, (strict) alveolo-palatal = 4.

(42) *HETEROSIB$DAC_{[+ant]}$ (*HETSIB): Adjacent $DAC_{[+ant]}$ consonants do not differ in the value of the feature [sibilant].

An uncontroversial markedness constraint, reflecting the fact that denti-alveolar place is unmarked for rhotics, is also involved here: *R[−ant]. It was established

at (20–24) above that the word–phrase correspondence constraint IDENTW-P{ʃKP}, preserving all but the least marked major place, outranks the marked cluster constraint *{ʃKPT}$_{[-\text{cont}]}${ʃKP}. Tableau (43) compares candidates for the input sibilant clusters, as (26c) *mateix so* /ʃ#s/ [s^{j}s^{j}] ∼ [s^{j}] 'same sound', (25a) *dos generals* /z#ʒ/ [ʒʒ] ∼ [ʒ] 'two generals'. Most of the lower-ranked constraints are omitted to simplify the presentation.

(43)

	/ʃs/	IDENTW-P{ʃKP}	*DISTDAC$_{[+\text{ant}]}$	IDENTW-P{ʃKPT}
(a)	i. ʃs		2!	
	ii. s^{j}s		1!	
	iii. ss	*!	0	*
	iv. ʃs^{j}		1!	*
	v. ☞ s^{j}s^{j}		0	*
(b)	/zʒ/			
	i. zʒ		2!	
	ii. z^{j}ʒ		1!	*
	iii. zz^{j}		1!	
	iv. ☞ ʒʒ		0	*

IDENTW-P{ʃKP} » *DISTDAC$_{[+\text{ant}]}$ » IDENTW-P{ʃKPT}

In (43a) IDENTW-P{ʃKP} preserves coda alveolo-palatal place, eliminating candidate (iii); then *DISTDAC$_{[+\text{ant}]}$ favours the most assimilated candidate, despite its violation of faithfulness to onset place (specifically [+anterior]). In (43b) likewise, the most assimilated candidate is preferred, despite neutralization of coda place. Except in formal styles of speech, the outcome of these clusters is in fact a non-geminate sibilant. That is, most styles are subject to a markedness constraint against geminate sibilant fricatives *GEMSib. (Kirchner (1998: 158) argues that the precise articulatory constraints for sibilant fricatives conflict with the tension and timing requirements of a geminate, with the effect that geminate sibilant fricatives are more marked than other geminates.) Assuming that high-ranked MAX constraints exclude the deletion interpretation of simplified geminates (i.e. a simplified output geminate here displays multiple correspondence with its two input consonants), simplified geminates follow from the constraints already given, with *GEMSib standing above IDENTW-P{ʃKPT}, and with IDENTOns[distr]DAC selecting the right candidate. Tableau (44) illustrates this alternative; constraints not relevant to this evaluation are omitted.

(44)

	/ʃ$_1$s$_2$/	IDW-P{ʃKP}	IDOn[distr]DAC	*DISTDAC$_{[+ant]}$	*GEMSib	IDW-P{ʃKPT}	UNIFORM
(a)	i. ʃs			2!			
	ii. s^{j}s			1!			
	iii. ss	*!		0	*	*	
	iv ʃs^j			1!		*	
	v s^{j}s^j			0	*!	*	
	vi. ʃ$_{12}$		*!			*	*
	vii. ☞ s^j$_{12}$					*	*
	viii s$_{12}$	*!				*	*
(b)	/z$_1$ʒ$_2$/						
	i. zʒ			2!			
	ii. z^{j}ʒ			1!		*	
	iii. zz^j		*!	1			
	iv. ʒʒ			0	*!	*	
	v. z$_{12}$	*!	*			*	*
	vi. z^j$_{12}$		*!			*	*
	vii. ☞ ʒ$_{12}$					*	*

IDENTW-P{ʃKP}, IDENTOnset[distr]DAC, *DISTDAC$_{[+ant]}$, *GEMSib » IDENTW-P{ʃKPT}, UNIFORMITY

In (44) the multiply corresponding alveolars (44a.viii, 44b.v) are eliminated by IDENTW-P{ʃKP} because the major place features of the alveolo-palatal inputs (C_1 in (44a), C_2 in (44b)) are missing. Of the remaining candidates in (44a) only (vii) is faithful to the [−distr] value of the input onset /s/, while in (44b) only (vii) is faithful to the [+distr] value of the input onset /ʒ/. UNIFORMITY is the correspondence constraint against fusion (45).

(45) UNIFORMITY: No element of S_2 has multiple correspondents in S_1 (McCarthy & Prince 1995: 123).

Tableau (46) evaluates candidates for the sibilant–rhotic clusters (*peix recuit* /ʒ#r/ [z^{j}ɹ^s] 're-cooked fish' (26c), and *dos reis* /z#r/ [ɹ^{s}ɹ^s] 'two kings' (27b)).

(46)

	/ʒr/	IDW-P{ʃKP}	*R[−ant]	*HETSib	*DISTDAC[+ant]	IDW-P{ʃKPT}
(a)	i. ʒr			*!	2	
	ii. zr	*!		*	0	*
	iii. zɹˢ	*!			0	*
	iv. ☞ zʲɹˢ				1	
	v. zʲrʲ		*!	*	0	*
	vi. zʲr			*!	1	
	vii. zʲɹʲˢ		*!		0	*
(b)	/zr/					
	i. zr			*!	0	
	ii. ☞ zɹˢ				0	
	iii. ☞ rr				0	

IDENTW-P{ʃKP}, *R[−ant], *HETSib »
*DISTDAC[+ant]

The high-ranking IDENTW-P{ʃKP} eliminates candidates (46a. ii, iii) that neutralize alveolo-palatal and denti-alveolar place, while *R[−ant] rules out palatalized rhotics as a repair to heterorganicity. *HETSib excludes the candidates with unassibilated rhotics, but in (46b) it allows [rr] as a realization of /zr/. For *dos reis* /z#r/ 'two kings', [rr] or [r] is undoubtedly the more widespread realization, even in formal styles. This would follow from a constraint against a marked consonant type, a sibilant rhotic, (*ɹˢ), standing below *HETSib in the ranking seen in (46). In this position the assibilated rhotic in the winning candidate of (46a) would survive, as the other candidates have already been eliminated.

6.5.5 *Conclusions and Summary*

This lengthy account of coronal contacts has revealed a series of anti-heterorganic marked cluster constraints of increasingly narrow focus, *DISTFront, *DISTDAC−DAC, and *DISTDAC[+ant]. The first applies in principle to all tongue-front consonants. In fact, it would be simpler overall to invoke a more general *DIST constraint not limited to tongue front consonants. Such a constraint would then deal with bilabial and labio-dental place assimilation in the same way. *DISTDAC−DAC focuses on the stronger necessity to avoid

(47)

Constraints in the resolution of coronal contacts		Indicative examples
IDENTW-P{ʃKP}, IDENTRhotic, IDENTW-P[cont], *DORSAL[+lat], *LABIAL[+lat]	cf. (24)	
»		
*{ʃKPT}$_{[-cont]}${ʃKP}, *R[−ant], IDENTOnset[back], *HETSib, IDENTOnset[distr]DAC		
»		nʲʒ ≻ nʒ, zʲɹˢ ≻ zʲrʲˢ, zʲɬ ≻ zʲl, ʒl, zʲɹˢ ≻ zʲrʲ, ʒʒ ≻ zʲzʲ, rʒ ≻ rʲʒ
*DISTDAC$_{[+ant]}$, *GEMSib		
»		zʲ ≻ zʲzʲ ≻ ʒz, zʲz
IDENTW-P{ʃKPT}	IDENTI-O{ʃKPT}, IDENTNas, IDENTLat	
»		lʒ ≻ lʲʒ
*DISTDAC		
»		zʲɬ ≻ ʒɬ
IDENT[distr]DAC$_{[+ant]}$, UNIFORMITY		
»		ʒð̠ ≻ zʲð
*{ʃ}		
»		nʲʒ ≻ ɲʒ
*DISTFront		
»		ʒð̠ ≻ ʒð
IDENTDent, IDENTAlv		

heterorganicity in consonants with a high degree of articulatory constraint (DAC). The third, *DISTDAC$_{[+ant]}$ restricts place differences between the consonant types with the highest degree of articulatory constraint, namely, sibilants and rhotic trills. Of course, these three members of the *HETERORG family are separated by several place-faithfulness constraints, together with a few markedness constraints. Notable are two faithfulness constraints that preserve, in particular, consonants with high DAC, namely, IDENT[distr]DAC$_{[+ant]}$ (31), protecting s/z, ʃ/ʒ, r, and a positional variant IDENTOnset[distr]DAC (38), protecting all DAC consonants s/z, ʃ/ʒ, r, ɲ, ʎ, ɬ, in onsets. I hope to have shown that the system, while complex, due to the range of different consonant types involved in contacts, is orderly and motivated, and can be accounted for by an OT approach which is phonetically informed. In (47) I set out a summary of the constraint hierarchy established in this section, which amplifies and links to the one established in (24).

TABLE 6.4 Realization of coda-C clusters in Majorcan (after Dols 1993, Recasens 1991b) Voiced consonants only are illustrated, except for coda s/z.

C_1 \ C_2		bil obs	bil nas	l-d fric	den obs	alv nas	alv lat	alv fric	alv trill	a-p fric	a-p nas	a-p lat	pal app	pal obs	vel obs
bil obs	1	bb													
	2				dd			d̪ːz		dʲːʒ				ɟɟ	gg
	3		mm	vv		nn	ɫɫ		rr		ɲɲ	ʎʎ	jj		
bil nas	1	mb	mm												
	2			ɱv	n̪d	nn	nt	nz	nr	nʲʒ	ɲɲ	nʲʎ	nʲj	nʲɟ	ɣ̃g
	3						ɫ̃ɫ					yʎ	jj		
	4						ɫɫ					ʎ̃ʎ			
l-d fric	1			vv											
	3	bb	mm		dd	nn	ɫɫ	d̪ːz	rr	dʲːʒ	ɲɲ	ʎʎ	jj	ɟɟ	gg
den obs/T	1				dd										
	2	bb						d̪ːz		dʲːʒ				ɟɟ	gg
	3		mm	vv		nn	ɫɫ		rr		ɲɲ	ʎʎ	jj		
alv nas/N	1					nn	nɫ	nz	nr						
	2	mb	mm	ɱv	n̪d		~			nʲʒ	ɲɲ	nʲʎ	nʲj	nʲɟ	ŋg
	3						ɫ̃ɫ					ɣ̃ʎ	jj		
	4						ɫɫ					ʎʎ			
alv lat	1	ɫβ	ɫm	ɫv		ɫn	ɫɫ	ɫz	ɫr						ɫɣ
	2				ɫ̪d					lʒ	lɲ	lʎ	lj	lɟ	
alv fric[–vc]	1	sp		sf											
	3				st̪			t̪ːs		tʲːʃ				sc	sk
alv fric	1	zβ	zm	zv		zn	zɫ				zɲ	zʎ	zj		
	2				zð̪									zɟ	zɣ
	3	ɾβ	ɾm	ɾv	ɾð	ɾn		d̪ːz	rr	dʲːʒ	ɾɲ			ɾɟ	ɾɣ
	4		mm			nn	ɫɫ				ɲɲ	ʎʎ	jj		
alv rhot	1	ɾβ	ɾm	ɾv	ɾð	ɾn	ɾɫ	ɾz	rr~r	ɾʒ	ɾɲ	ɾʎ	ɾj	ɾɟ	ɾɣ
	3						ɫɫ					ʎʎ			
a-p fric	3	jβ	jm	jv	jð	jn	jɫ	jz	jr	jʒ	jɲ	jʎ	jj	jɟ	jɣ
	4									dʲːʒ					
a-p nas	1										ɲɲ				
	2														
	3	jmb	jmm	jɱv	j̪nd	jnn	j̪nɫ	j̪nz	j̪nr	j̪nʲʒ	j̪ɲɲ	j̪nʲʎ	j̪nʲj	j̪nʲɟ	j̪ŋg
	4						jɫ̃ɫ					jl̃ʎ	jj		
a-p lat	1	ʎβ	ʎm	ʎv	ʎd	ʎn	ʎl	ʎz	ʎr	ʎʒ	ʎɲ	ʎʎ	ʎj	ʎɟ	ʎɣ
	2														
	3						jl						jj		
pal/vel obs	1														
	2	bb			dd			d̪ːz		dʲːʒ				ɟɟ	gg
	3		mm	vv		nn	lɫ		rr		ɲɲ	ʎʎ	jj		

Within each cell, row 1 = realizations faithful to input, row 2 = coarticulation/place assimilation, row 3 = realizations involving manner assimilation, row 4-realizations involving further manner assimilation. The coronal+coronal contacts are those within the bold box. The greyed cells correspond to clusters not found within morphemes, for principled reasons. Abbreviations: bil = bilabial, l-d = labiodental, den = dental, alv = alveolar, a-p = alveolo-palatal, pal = palatal, vel = velar; obs = obstruent, fric = fricative, nas = nasal, lat = lateral, rhot = rhotic, app = approximant.

6.6 CONSONANTAL CONTACT IN MAJORCAN: PLACE AND MANNER

One of the most notable differences between the phonology of Balearic and continental Catalan (more exactly, between Majorca and Minorca on the one hand and the continent and Ibiza on the other) lies in the realization of consonant clusters. Majorcan Catalan displays radical assimilation of place, and notable assimilation of manner, in coda consonants. In Minorca there are some differences of detail, and in general assimilation is somewhat less radical, in formal styles, at least. Only the Majorcan data are examined here. Table 6.4 presents the results of contact for Majorcan in a form parallel to that in Table 6.1 with respect to continental Catalan.

In Table 6.4 the row corresponding to alveolar fricatives is split in two, with one row for voiceless fricatives, the other for voiced. This is because [z] before voiced consonants may be realized as a rhotic (tap [ɾ] or approximant [ɹ]), in informal styles in Majorca, as indicated in the table, and this rhotic is also subject to optional manner assimilation to a following nasal or liquid. Table 6.4 also contains a column for palatal obstruents since these are common variants of the dorsal phonemes in Majorca (see §2.1.1). The greyed-in area is much larger in Table 6.4 than in Table 6.1. Within morphemes in Majorcan there are no heterorganic place clusters among stops and nasals. The marked cluster constraints are highly ranked in Majorcan, and the morpheme-internal heterorganic clusters found in other varieties have been lexically restructured, e.g. *premsa* Majorcan /pɾens+a/ 'press'.

Whereas in continental Catalan the result of major place contact is largely driven by the ranking IDENTW-P{ʃKP},*{ʃKPT}$_{[-\text{cont}]}${ʃ} » IDENTW-P {ʃKPT}—i.e., by faithfulness to non-denti-alveolar place plus the banning of heterorganic C–alveolo-palatal clusters, together outranking general place faithfulness—in Majorcan only faithfulness to alveolo-palatal place (IDENTW-P{ʃ}) outranks two general marked cluster constraints, first *{ʃKP}$_{[+\text{cons}]}${ʃKPT}, which bars all heterorganic place except for denti-alveolars or glides before a consonant, and secondly *{ʃKPT}$_{[-\text{cont}]}${ʃKPT}, which bars any heterorganic place with a non-continuant coda (with the effect that the only heterorganic clusters are those with glides, liquids, or alveolar fricatives in codas). These constraints interact in some cases with those involving manner differences—stops, nasals, laterals, rhotics; manner assimilation is largely driven by the Syllable Contact Law (SYLCON), as above. Other interesting features of Majorcan consonant contact are affrication as a 'repair' to *GEMINATESibilant violations (*dos senyors* [dotːsəˈɲos] 'two gentlemen'), the fact that coda /f/ is treated exactly the same as a stop (*agaf pomes* [əˈjapˈpoməs] 'I pick up apples'), and 'compensatory diphthongization' (*anys difícils* /ˈaɲ+z#diˈfisil+z/ [ˈajn̪di ˈfisils] 'difficult years', where /ɲ/ is realized as a sequence [jN]; see Mascaró (1986b)). These phenomena are included in the account below, alongside place assimilation.

I look first at Majorcan clusters involving manner in consonantal sonorants (ɾ/r, n, l, ʎ). (Majorcan generally has [ɾ] as the realization of the rhotic /R/ in codas, unlike the pattern discussed above from continental eastern Catalan which has [r] in most codas.) As Table 6.4 shows, sonorant clusters may all be resolved faithfully as regards manner, but assimilation of rhotics and nasals to laterals is possible. Some examples are given in (48). (I return subsequently to the resolution of marked place in coda sonorants.)

(48) *un líquid* [ɬɬ] ∼ [ɬ̃ɬ] ∼ [nɬ] 'a liquid' *molt ric* [ɬr] *[rr] 'very rich'
un lleó [ʎʎ] ∼ [nʲʎ] 'a lion' *molt notable* [ɬn] *[nn] 'very remarkable'
mur llest [ʎʎ] ∼ [ɾʎ] 'ready wall' *mur nou* [ɾn] *[nn] 'new wall'

What constraint interaction gives rise to the non-faithful options? Highly ranked in Majorcan (probably undominated) is faithfulness to word-initial onsets IDENTW-PCues(IDCues). The markedness constraint provoking assimilation of manner is, I suggest, one penalizing sonorant clusters of conflicting manner, as in (49).

(49) *HETEROManner$_{[+son]}$ (*HETMan): Adjacent consonantal sonorants do not differ in the value of [lateral], [nasal], or [rhotic].

The constraint *HETMan interacts with sonorant faithfulness constraints, IDENTRhotic, IDENT[+lateral], IDENTNasal, IDENT[±lateral]. Of the two versions of *un líquid* 'a liquid' with lateral assimilation [ɬɬ] ∼ [ɬ̃ɬ] the second has a nasalized lateral which retains faithfulness to [nasal]. Consequently there must be a markedness constraint that can prefer [ɬɬ], which I take merely to be a constraint against nasalized laterals as complex segments: *ɬ̃. SYLCON is the Syllable Contact constraint discussed in §3.1. Tableau (50) illustrates these constraints with respect to the first alternatives given in (48).

(50)

	/nɬ/	IDCues	SYLCON	ID[+lat]	*ɬ̃	*HETMan	IDRhot	IDLat	IDNas
(a)	i. nɬ		*!			*			
	ii. ☞ɬɬ							*	*
	iii. nn	*!		*				*	
	iv. ɬ̃ɬ				*!	*		*	
(b)	/RʎʎR/								
	i. ɾʎ					*!			
	ii. ☞ʎʎ						*	*	
	iii. rr	*!		*			*	*	

(c)

ɫn	IDCues	SYLCON	ID[+lat]	*ɫ̃	*HETMan	IDRhot	IDLat	IDNas
i. ☞ ɫn					*			
ii. nn			*!				*	
iii. ɫɫ	*!						*	*

(Majorcan) IDENTCues, SYLCON, IDENT[+lat], *ɫ̃ » *HETMan » IDENTRhot, IDENTLat, IDENTNas

Candidate (50a.i) violates SYLCON because a coda nasal is followed by a more sonorant liquid; it also violates the *HETMan constraint. In the variety which retains [nɫ] in *un líquid*, SYLCON and *HETMan rank below IDENTLat or IDENT-Nasal. In the case of (50b), (i)'s violation of *HETMan is fatal, and [ʎʎ] with an unfaithful coda is preferred. In (50c) /ɫn/ remains unassimilated because ID[+lat] outranks *HETMan.

In Majorcan all non-sibilant obstruents (p/b, t/d, k/g (or c/ɟ) and f/v) assimilate to the place and manner of following onsets, giving rise to geminates. The only apparent exception is a stop preceding a sibilant fricative, where the outcome is a sibilant affricate, for example, *trec sucre* [ˈtrɛˈt̪ːsukɾə] 'I take out sugar'—no real exception, because sibilant affricates replace geminate sibilant fricatives in any case. Non-dorsal nasals (m, n) also assimilate place, and optionally, as seen above, assimilate to laterals in manner. Dorsal nasals (ɲ, and n^j or ŋ from underlying Nc/Nɟ or Nk/Ng) retain some cues to underlying palatality, as is the case with the alveolo-palatal lateral (ʎ) and the alveolo-palatal fricatives ʃ/ʒ. The data below is taken largely from Dols (1993: 93–7).

(51) Underlying bilabial–C clusters in Majorcan

(a) obstruents

rep massa [ˈrəmˈmasə] 'receives too much'
rep forasters [ˈrəfforəsˈtes] 'receives strangers'
rep turistes [ˈrəttuˈɾistəs] 'receives tourists'
rep notícies [ˈrənnoˈtisis] 'receives news'
rep làmines [rəɫˈɫaminəs] 'receives plates'
rep sucre [ˈrəˈtːsukɾə] 'receives sugar'
rep roses [ˈrərˈrɔzəs] 'receives roses'
rep xots [ˈrəˈtːʃɔts] 'receives lambs'
rep llàgrimes [ˈrəʎˈʎaɣɾiməs] 'receives tears'
rep iots [ˈrəjˈjɔts] 'receives yachts'
rep coses [ˈrəkˈkɔzəs] 'receives things'

(b) nasals

som vint [ˈsomˈvin̪t] 'there are twenty of us'
som tres [ˈson̪ˈtɾəs] 'there are three of us'
som nedador [ˈsonneðəˈðo] 'I am a swimmer'
som legal [ˈsoɫɫəjaɫ] ~ [ˈsonɫəjaɫ] 'I am legal'

som cinc [ˈsonˈsinʲc] ‘there are five of us’
collim nyores [koˈiɲɲɔɾəs] ‘we harvest chillies’
som lleial [ˈsoʎʎeˈaɫ] ∼ [ˈsonʲʎeˈaɫ] ‘I am loyal’
som hienes [ˈsonʲˈjenəs] ‘we are hyenas’
som quinze [ˈsonʲˈcinzə] ‘there are fifteen of us’
som quatre [ˈsoŋˈkwatɾə] ‘there are four of us’

(52) Underlying labio-dental–C clusters in Majorcan
agaf plats [əˈʝapˈplats] ‘I pick up plates’
agaf martells [əˈʝamˈmərˈteʎs] ‘I pick up hammers’
agaf trossos [əˈʝatˈtɾɔsos] ‘I pick up pieces’
agaf nins [əˈʝanˈnins] ‘I pick up children’
agaf líquid [əˈʝaɫˈɫicit] ‘I pick up liquid’
agaf sucre [əˈʝaˈtːsukɾə] ‘I pick up sugar’
agaf roques [əˈʝarˈɾɔcəs] ‘I pick up rocks’
agaf llana [əˈʝaʎˈʎanə] ‘I pick up wool’
agaf iogurts [əˈʝajjoˈɣuɾs] ‘I pick up yogurts’
agaf quatre [əˈʝakˈkwatɾə] ‘I pick up four’

(53) Underlying denti-alveolar non-strident obstruents/nasal–C clusters in Majorcan
(a) obstruents
tot ple [ˈtopˈplə] ‘completely full’
tot meu [ˈtomˈmew] ‘all mine’
tot falla [ˈtofˈfaʎə] ‘everything fails’
tot net [ˈtonˈnət] ‘completely clean’
tot líquid [ˈtoɫˈɫicit] ‘completely liquid’
tot roda [ˈtorˈɾɔðə] ‘everything runs’
tot lluent [ˈtoʎʎuˈent] ‘all shining’
tot cou [ˈtokˈkɔw] ‘everything stings’
(b) nasals
són vint [ˈsomˈviɲ̪t] ‘there are twenty of them’
són tres [ˈsoɲ̪ˈtɾəs] ‘there are three of them’
són nedadors [ˈsonneðəˈðos] ‘they are swimmers’
són legals [ˈsoɫɫəˈʝaɫs] ∼ [ˈsonɫəˈʝaɫs] ‘they are legal’
són cinc [ˈsonˈsinʲc] ‘there are five of them’
cullen nyores [ˈkujəɲˈɲɔɾəs] ‘they harvest chillies’
són lleials [ˈsoʎʎeˈals] ∼ [ˈsonʲʎeˈaɫs] ‘they are loyal’
són hienes [ˈsonʲˈjenəs] ‘they are hyenas’
són quinze [ˈsonʲˈcinzə] ‘there are fifteen of them’
són quatre [ˈsoŋˈkwatɾə] ‘there are four of them’

(54) Underlying palatal/velar non-strident obstruent–C clusters in Majorcan
poc pa [ˈpɔpˈpa] ‘not much bread’
poc millor [ˈpɔmmiˈʎo] ‘not much better’
poc vi [ˈpɔvˈvi] ‘not much wine’
poc temps [ˈpɔtˈtens] ‘not much time’

poc nou [ˈpɔnˈnɔw] 'not much new'
poc legal [ˈpɔɬɬəˈʝal] 'not very legal'
poc ros [ˈpɔrˈros] 'not very fair-haired'
poc sucre [ˈpɔˈt̪s͡ukɾə] 'not much sugar'
poc generós [ˈpɔdʲːʒənəˈɾos] 'not very generous'
poc lliure [ˈpɔʎˈʎiwɾə] 'not very free'
poc iogurt [ˈpɔjˈjoˈɣuɾt] 'not much yogurt'

The assimilations in (51–4) display the effects of three quite general constraints. The first is SYLCON, which penalizes a coda that is less sonorant than a following onset. The second is a marked cluster constraint much more far-reaching than the constraint $*\{\text{ʃKPT}\}_{[-\text{cont}]}\{\text{ʃ}\}$ that is observed in continental Catalan (20). The Majorcan constraint is the one in (55).

(55) $*\{\text{ʃKPT}\}_{[-\text{cont}]}\{\text{ʃKPT}\}$: No heterorganic clusters when C_1 is [−continuant].

The third is a place-markedness constraint penalizing consonant place other than denti-alveolar: $*\{\text{ʃKP}\}_{[+\text{cons}]}$. In Majorcan the only heterorganic clusters are those in which the first consonant is a glide (j/w) or a denti-alveolar (ɾ/ɾ, ɬ, s/z). Labio-dental fricatives do not even occur before other labials (i.e. no *fp, *vb, *vm), nor do alveolo-palatal fricatives occur before other alveolo-palatals (*ʒʎ, *ʒɲ). Place in onsets and in utterance-final position is, of course, protected by IDENTCues. And, as we shall see, alveolo-palatal place is not completely obliterated in codas. With high-ranking IDENT{ʃ} and lower $*\{\text{ʃKP}\}_{[+\text{cons}]}$, alveolo-palatal place can only be realized in a palatal glide [j]. (Alveolo-palatal laterals and nasals are better preserved in Majorcan as a consequence of higher-placed IDENTLateral and IDENTNasal, to be discussed below.) The correspondence alignment constraint UNIFORMσ (§3.1.1 (15)) has a role too in distinguishing candidates. The INTEGRITY constraint (56) from correspondence theory is also relevant; the formulation is that of McCarthy & Prince (1995: 124), though in (57) its effects mirror those of UNIFORMσ, and it is not displayed separately in the tableau.

(56) INTEGRITY: 'No breaking'. No element of S_1 has multiple correspondents in S_2.

6.6.1 *Labial coda inputs*

In tableau (57) I display some candidates for labial stops followed by obstruents corresponding to the following examples from (51) *rep forasters* [ˈrəfforəsˈtes] 'receives strangers', *rep turistes* [ˈrəttuˈɾistəs] 'receives tourists', *rep sucre* [ˈrəˈt̪s͡ukɾə] 'receives sugar', *rep xots* [ˈrəˈtːʃ͡ɔts] 'receives lambs'. Here and subsequently candidates violating IDENTCues (essentially, onset faithfulness) are omitted.

In (57a) candidate (iii) avoids the SYLCON violation seen in candidate (i) by affrication+gemination, but this causes a violation of Word syllable alignment UNIFORMσ, since coda /p/ is now also associated with the onset of the following

syllable. Candidate (ii) survives despite its violation of IDENT[continuant]. In (57b) candidate (i) has a straightforward marked cluster violation, to which an assimilated candidate (ii) is preferable. Vocalization of /p/ to [w] in coda is a conceivable repair to a heterorganic stop cluster. In the approach offered here it is excluded (iii) by high-ranking IDENT[−back]. Candidate (iii) does also have an IDENT[cont] violation, but vocalization of /p/b/ is not acceptable even when continuance is not at issue. However, excluding vocalization by IDENT[cons] would be mistaken, as alveolo-palatal obstruents are indeed realized as (palatal) glides. In (57c) the place+manner assimilated candidate (iii) is ruled out by *GEMSib, leaving the homorganic affricate the winner. A conceivable *[p.ps] is not considered here; it would save the SYLCON violation, but would still fall foul of the marked cluster constraint. (It would also violate MINIMUM SONORITY DISTANCE, since a stop and a fricative are not sufficiently distant on the sonority scale to make an acceptable onset.) An affricate [t͡s] (homorganic) is interpreted as a single segment with respect to MSD. The pattern of (57c) is repeated in (57d).

(57)

	/p₁f₂/	ID[−bk]	*GEMSib	SYLCON	UNIσ	*{ʃKPT}[−cont].{ʃKPT}	*{ʃKP}	ID[cont]	IDW-P{ʃKPT}
(a)	i. pf			*!			**		
	ii. ☞ ff						*	*	
	iii. $p_1.p_1f_2$(= p͡ːf)				*!		*		
(b)	/pt/								
	i. pt					*!	*		
	ii. ☞ tt								*
	iii. wt	*!						*	
(c)	/p₁s₂/								
	i. p.s			*!		*	*		
	ii. ☞ $t_1.t_1s_2$ (= tːs)				*				*
	iii. s.s		*!					*	*
	iv. ws	*!							*
(d)	/p₁ʃ₂/								
	i. pʃ			*!		*	**		
	ii.☞ $tʲ_1.tʲ_1ʃ_2$ (= tʲːʃ)				*		*	*	*

(Majorcan) IDENT[−bk], *GEMSib, SYLCON » UNIFORMσ » *{ʃKPT}[−cont]{ʃKPT} » *{ʃKP}, IDENT[cont], IDENTW-P{ʃKPT}

In (58) bilabials before sonorants are considered, corresponding to these examples from (51): *rep massa* [ˈrəmˈmasə] ‘receives too much’, *rep notícies* [ˈrənnoˈtisis] ‘receives news’, *rep làmines* [rəɫˈɫaminəs] ‘receives engravings’, *rep roses* [ˈrərˈrɔzəs] ‘receives roses’, *rep iots* [ˈrəjˈjɔts] ‘receives yachts’. (As before, the inputs are represented voiced, since AGREEVoice is undominated.) IDENT[−back] violations are omitted, as are violations of *HETEROManner[+son], which, as seen in (50), outranks IDENTNas, IDENTLat and IDENTRhot.

(58)

	/bm/	SYLCON	UNIσ	IDNas	IDLat	IDRhot	*{ʃKPT}[−cont]{ʃKPT}	*{ʃKP}	ID[cont]	IDW-P{ʃKPT}
(a)	i. bm	*!						**		
	ii. ☞ mm			*				*		
(b)	/bn/									
	i. bn	*!						*		
	ii. mn			*			*!	*		
	iii. ☞ nn			*						*
(c)	/b_1.$ɫ_2$/									
	i. b.ɫ	*!					*	*		
	ii. .βɫ		*!				*	*	*	
	iii. b_1.$b_1ɫ_2$		*!				*	*		
	iv. ☞ ɫɫ				*				*	*
(d)	/br/									
	i. br	*!					*	*		
	ii. dr	*!								*
	iii. ☞ rr					*			*	*
(e)	/bj/									
	i. bj	*!					*	*		*
	ii. ☞ jj								*	

(Majorcan) SYLCON » UNIFORMσ » IDENTNas, IDENTLat, IDENTRhot, *{ʃKPT}[−cont]{ʃKPT} » *{ʃKP}, IDENT[cont], IDENTW-P{ʃKPT}

In (58) SylCon knocks out all the obstruent–sonorant candidates, in (58b) *{ʃKPT}$_{[-\text{cont}]}${ʃKPT} picks among the survivors. Candidates such as [ɾn] or [łn] for (58b) are not illustrated. It is these that *HetMan excludes. In (58c) the Alignment correspondence constraint Uniformσ discounts some resyllabified repairs of /bł/ that preserve labial place. Such resyllabifications of (58d) and (58e) would fail equally, of course.

6.6.2 *Alveolar coda inputs*

I turn next to some cases with coda alveolar sibilants in the input. Alveolar sibilants are retained before voiceless onsets, and are retained, except in informal styles (where [ɾ] or manner assimilation occurs), before most voiced sounds. Before /ɾ/, /z/ is realized as [r], i.e. with manner assimilation, as generally on the continent, I assume here for the same reasons. Sibilant sequences, voiceless and voiced, become affricates, however. Some examples follow (59).

(59) (a) *es paper* [əspəˈpe] 'the paper' *es monjo* [əzˈmɔnʲʒo] 'the monk'
es fetge [əsˈfedʲːʒə] 'the liver' *es iogurt* [əzjoˈɣurt] 'the yogurt'
(b) *ses sabates* [sə͡t̪səˈβatəs] 'the shoes' *es gelat* [ə͡dʲːʒəˈłat] 'the ice cream'

Such sequences of alveolar sibilant–consonant do not violate the marked cluster constraint *{ʃKPT}$_{[-\text{cont}]}${ʃKPT} as s/z are continuants. But some of them, for example /zm/, do violate SylCon. I suggest that IdentSibilant is the constraint that protects them. In this respect, the variety that replaces coda [z] with [ɾ] or assimilation illustrates the 'emergence of the unmarked', inasmuch as the tap approximant [ɾ] forms a less marked coda than a sibilant. Thus in informal style suppression of IdentSibilant allows SylCon to favour variants with coda sonorants (though there must be other reasons to account for *es dia* [əɾˈðiə] 'the day', since [z.ð] is an acceptable syllable contact). Other than IdentSib, the constraints relevant here are *GemSib, mentioned above in §6.5 and (46), penalizing geminate sibilant fricatives, and *HetSib (43), which disfavours contiguous sibilants of different place.

The 'repair' of geminate sibilants as geminate affricates in Majorcan is often seen as rather unusual. However, Kirchner's 'effort-based' theory (1998) has this precise result: 'the geminate fricative . . . emerges as more effortful than . . . the geminate affricate' (1998: 158). The *GemSib constraint is thus one that derives from an effort-based markedness scale, in which geminate strident fricatives stand towards or at the most marked end. Tableau (60) displays several clusters from the examples in (59).

(60)

	/sp/	*GEMSib	*HETSib	IDSib	SYLCON	UNIσ	*{ʃKP}	ID[cont]	IDW-P{ʃKPT}
(a)	i. ☞ sp						*		
	ii. fp			*!			**		*
	iii. pp			*!			*	*	*
(b)	/zm/								
	i. ☞ zm				*		*		
	ii. mm			*!			*	*	*
	iii. ɾm			*!			*		
(c)	/s_1.s_2/								
	i. s.s	*!							
	ii. ɾs			*!					
	iii. ☞ $\widehat{ɫ_1.ɫ_1}s_2$					*		*	
(d)	/$z_1ʒ_2$/								
	i. zʒ		*!				*		
	ii. ʒʒ	*!					*		
	iii. ☞ $\widehat{dʲ_1.dʲ_1}ʒ_2$					*	*	*	

(Majorcan) *GEMSib, *HETEROSib, IDENTSib » SYLCON » UNIFORMσ » [IDENTNas, IDENTLat, IDENTRhot, *{ʃKPT}[–cont]{ʃKPT} »] *{ʃKP}, IDENT[cont], IDENTW-P{ʃKPT}

6.6.3 *Alveolo-palatal coda inputs*

The resolution of alveolo-palatal–C codas in Majorcan raises interesting issues, because the three types of alveolo-palatal involved, sibilant ʃ/ʒ, nasal ɲ, and lateral ʎ, are resolved differently, as seen in (61a, b):

(61) (a) *mateix dia* /məˈteʃ#ˈdiə/ [məˈtejˈðiə] ‘same day’
mateix senyor [məˈtejsəˈɲo] ‘same gentleman’
mateix gènere [məˈtejˈʒɛnərə] ‘same kind’
(b) *any passat* /ˈaɲ#paˈsad/ [ˈajmpəˈsat] ‘last year’
company jove /komˈpaɲ#ˈʒove/ [komˈpajnʲˈʒovə] ‘young companion’
ball difícil /ˈbaʎ#diˈfisiɫ/ [ˈbaʎdiˈfisiɫ] ‘difficult dance’

(c) *tan jove* /ˈtan#ˈʒove/ [ˈtanʲˈʒovə] 'so young'
cranc jove /ˈkɾaNc/ [ˈkɾanʲc] + /ˈʒovə/ [ˈkɾajnʲˈʒovə] 'young crab'
menj prunes /ˈməNʒ/ [ˈmənʲʒ] +/ˈpɾunəs/ [ˈməjmˈpɾunəs] 'I eat plums'

The alveolo-palatal fricative is realized as a palatal glide preconsonantally, with loss of sibilance and consonantal place. The alveolo-palatal nasal is realized by 'breaking' or 'diphthongization' (violating INTEGRITY); a palatal glide precedes a homorganic coda nasal. The alveolo-palatal lateral is realized faithfully. In every case palatality is retained in some form. This follows from a high ranking of IDENTPalatal, where 'Palatal' is not restricted to consonantal place but includes vocalic place (though excluding palatalized alveolar place). Notice that IDENTPalatal is not identical to the IDENT{ʃ} element of IDENT{ʃKP} seen in (9) and subsequently in the account of continental clusters. IDENT{ʃ} is satisfied by consonants which are [−ant, −high], while IDENTPalatal is satisfied by consonants and glides which are [−ant, +distr, −back]. That is to say, Majorcan insists on a more strongly palatal realization of its alveolo-palatals (and palatals) than continental Catalan. IDENTPalatal interacts with the previously mentioned markedness constraint *{ʃ} which penalizes strictly consonantal alveolo-palatals [−ant, +distr, −high] (not palatalized alveolars or palatal glides). The difference between /ɲ/ realized [jN] and /ʎ/ realized [ʎ] (not *[jɫ]) suggests different versions of INTEGRITY are operating, one variety being restricted to non-nasals. I label this version INTEGRITY[−nas] (INTEG[−N]). It may be that nasals are treated differently because the nasal part of the split variant (jN) can assimilate fully to following consonant place, whereas the consonantal element of the other potential split alveolo-palatals, (*jl, *jz), must necessarily disagree in place with certain following consonants. The issue invites further investigation. Sibilant input coda clusters exemplified in (61a) are illustrated in (62).

(62) (a)

/ʒ$_1$d$_2$/	*GEMSIB	*HETSIB	IDPAL	INTEG [−N]	*{ʃ}	IDSIB	SYLCON	IDNAS	INTEG
i. ʒð					*!				
ii. ☞ jð						*			
iii. zð			*!						
iv. j$_1$z$_1$.ð$_2$				*!					
v. z^{j}ð			*!						
vi. dd			*!			*			

(b)

/ʃ$_1$s$_2$/	*GEMSIB	*HETSIB	IDPAL	INTEG [–N]	*{ʃ}	IDSIB	SYLCON	IDNAS	INTEG
i. ʃs		*!			*!				
ii. ☞ js									
iii. ss	*!		*						
iv. t̪$_1$.t̪$_1$s$_2$= t̪ːs			*!						
v. j$_1$t̪$_1$s$_2$				*!		*			*
vi. s^{j}s		*!	*						

(c)

/ʒ$_1$ʒ$_2$/									
i. ʒʒ	*!								
ii. ☞ jʒ					*	*			
iii. d^{j}d^j$_1$ːʒ$_2$= d^{j}ːʒ				*!	*				*

(Majorcan) *GEMSIB, *HETEROSIB, IDENTPAL, INTEGRITY[–N] » *{ʃ} » IDENTSIB [» SYLCON » UNIFORMσ » IDENTNAS, INTEGRITY, IDENTRHOT, *{ʃKPT}$_{[cont]}${ʃKPT} » *{ʃKP}, IDENT[cont], IDENTW-P{ʃKPT}]

Candidate (i) in (62a) would be the winner in mainland Catalan. In Majorca palatal consonants are avoided in Codas (by *{ʃ}), but palatality of some sort must be preserved (by IDENTPal), so (iii), (v), and (vi) lose out; splitting is not admissible (iv) except in nasals. Examples (62b, c) illustrate the strong constraints on sibilant clusters including geminates in Majorcan. As in (62a) the alveolo-palatal codas have to be realized as a front glide before a consonant.

Tableau (63) looks at candidates for the examples in (61b), involving input alveolo-palatal sonorants. Irrelevant constraints are omitted.

(63)

(a)

/ɲ$_1$p$_2$/	IDPAL	ID[+lat]	INTEG[–N]	*{ʃ}	IDNAS	*{ʃKPT}$_{[-cont]}${ʃKPT}	ID[cont]	INTEG
i. ɲp				*!				
ii. mp	*!							
iii. jp					*!		*	
iv. ☞ j$_1$m$_1$.p$_2$								*
v. j$_1$n$_1$.p$_2$						*!		*
vi. pp	*!				*			
vii. ĵp							*!	

(b) /ʎ₁d₂/	IDPal	ID[+lat]	INTEG[×N]	*{ʃ}	IDNas	*{ʃKPT}[–cont]{ʃKPT}	ID[cont]	INTEG
i. ☞ ʎd^gb				*				
ii. ʎd				*		*!		
iii. ɫ̪d	*!							
iv. jð		*!					*	
v. j₁ɫ₁d₂			*!					
vi. lʲd	*!							
(c) /ɲ₁ʒ₂/								
i. ɲʒ				**!				
ii. jʒ				*	*!			
iii. nʲʒ	*!			*				
iv. ☞ j₁nʲ₂ʒ				*				*
v. j₁n₂ʒ				*		*!		*

(Majorcan) IDENTPal, IDENT[+lat], INTEGRITY[–N] » *{ʃ} » IDENTNas, *{ʃKPT}[×cont]{ʃKPT}, IDENT[cont] » INTEGRITY

Because INTEGRITY[−N] does not penalize splitting /ɲ/ into [jN], candidates (iv) and (v) of (63a) survive to be passed down for further evaluation by *{ʃKPT}[−cont]{ʃKPT}, by which the more assimilated wins. Candidate (vii) with a nasalized glide also survives the higher constraints, but falls to IDENT [cont]. I assume [ʎd] is realized with gestural blending in Majorca as on the continent and have represented thus the winner (i) in (63b); but note that even unblended [ʎd] would win over the remaining candidates, which fail to be faithful to Palatal, [+lateral] or INTEGRITY[−N]. In (63c) note that the high ranking of *{ʃ}, compared to its position in continental Catalan, means that in Majorca the faithful and homorganic [ɲʒ] is less acceptable than the candidate (iv) that violates INTEGRITY and loses to (i) on DISTANCE constraints.

I turn finally to the clusters of (61c), where I ignore for the moment the details of cluster reduction that have the effect, in Majorcan, of eliminating all but the first member of a coda cluster when a consonant follows. Only candidates with such deletion are considered here; the word-form inputs are taken to be the forms in utterance-final position. The result makes more sense if we assume that [nʲc] in *cranc* [ˈkranʲc] represents gestural blending in which the contact spreads from palatalized alveolar to palatal (or perhaps only part of this area), so that the cluster as a whole is 'Palatal' as regards the IDENTPalatal constraint.

Since the coda consonant in (64a) is not underlyingly palatal, it survives with the best assimilation that passes *{ʃ} and INTEGRITY, namely palatalized alveolar. Comparison (64b) has two candidates (ii) and (iii) which differ in corres-

pondence relations, not in phonetic substance. Candidate (ii), where the elements are one-to-one correspondence (passing INTEGRITY), incurs two IDENTNas violations because input nasality is expressed in the second correspondent and not in the first. Candidate (iii) spreads the correspondence so each correspondent matches some aspects of the input, though not in a one-to-one way. Competition (64c) has two 'splitting' winners that do not differ in phonetic substance, so the choice between them is immaterial.

(64)

	/$n_1ʒ_2$/	IDPal	ID[+lat]	INTEG[–N]	*{ʃ}	IDNas	*{ʃKPT}$_{[-cont]}${ʃKPT}	ID[cont]	INTEG
(a)	i. nʒ				*		*!		
	ii. ɲʒ				**!				
	iii. ☞ $n^jʒ$				*				
	iv. $j_1n^j{}_2ʒ$				*				*!
(b)	/$n^j{}_1c_2{}^{gb}ʒ_3$/								
	i. $n^j{}_{12}ʒ$	*!			*				
	ii. $j_1n^j{}_2ʒ$				*	*!*			
	iii. ☞$j_{12}n^j{}_{12}ʒ$				*				*
(c)	/$n^j{}_1ʒ_2p_3$/								
	i. mp	*!							
	ii j_1m_2p	*!				*		*	
	iii. ☞ $j_{12}m_{12}p$								*
	iv. ☞ j_1m_1p								*

(Majorcan) IDENTPal, ID[+lat], INTEGRITY[–N] » *{ʃ} » IDENTNas, *{ʃKPT}$_{[-cont]}${ʃKPT}, IDENT[cont] » INTEGRITY

7

CLUSTER REDUCTION

7.1 INTRODUCTION

One of the phenomena of Catalan phonology that has given rise to some discussion in the phonological literature is 'homorganic stop deletion', the process by which, for example, *pont* /pɔnt/ 'bridge' is pronounced [pɔn] in most of Catalonia and in Ibiza. This process is usually linked to a more general one of cluster reduction affecting complex codas. In many varieties, a preconsonantal coda is generally limited to two consonants; thus *parts baixes* /parɫ+z baʃ+a+z/ 'lower parts' is pronounced [ˈparz.ˈβaʃəs], a reduction process that has yet more far-reaching effects in Majorcan. In Majorcan, with a limited exception, no more than one consonant can stand in a pre-consonantal coda, so in that variety *parts baixes* is pronounced [ˈpar ˈbaʃəs]. Dols (2002) gives a survey of clusters and cluster reduction in Catalan. Within Generative Phonology Catalan homorganic stop deletion has been treated by Lleó (1970: 25–9), Wheeler (1975/9: 280–87), Mascaró (1983: 100–11; 1985; 1989), Iverson (1993a, 1993b), Morales (1995), and Bonet & Lloret (1998: 108–14). These authors were largely concerned with correctly identifying the limits of the phenomenon, and formulating deletion rules accordingly. Majorcan clusters in particular are the focus in Dols (1993; 2001). Optimality perspectives on cluster deletion were introduced to the Catalan data by Colina (1995) and by Jiménez (1996a; 1999: 189–242). Jiménez's account provides data (especially Valencian) and inspiration for the treatment offered here, though there are significant differences. As usual, Recasens provides comprehensive data (1993: 167–9, 196–201, and esp. 1991b: 181–351 for dialect variation). Majorcan data originally assembled by Bibiloni are set out in Dols (1993), and given an Optimality account in Dols (2000: 311–48). The works of Dols (1993; 2000) are the source of most of the Majorcan data presented here. I use the general terminology 'cluster reduction' so as not to prejudice the interpretation of the phenomenon. From the perspective of OT correspondence theory, reduction could reflect deletion in the strict sense (with MAX violation), or alternatively, fusion (with some IDENT violation, along with UNIFORMITY violation).

In the account that follows I distinguish two pronunciation styles for the Catalan of Catalonia. The more formal variety, which is not excessively formal, corresponds to the description generally given by Recasens (1991b; 1993); not

surprisingly, it displays less reduction of clusters than the less formal variety, which is the variety treated by Jiménez (1999) in comparison with his own Valencian.

7.2 CLUSTER REDUCTION IN STEMS

In (1) I present the basic types of the words concerned. The orthography indicates the nature of the underlying coda clusters pretty well, and the underlying sequences are also clearly revealed before vowel-initial suffixes, as illustrated in (2). In (2) only central Catalan pronunciation is given; the pronunciation of the other varieties does not differ in relevant respects.

(1)	Final clusters bare stems	Catalonia more formal	Catalonia less formal	Valencia	Majorca
	(a) *arc* 'arch'	ark	ark	aɾk	aɾc
	serp 'snake'	serp	serp	seɾp	seɾp
	golf 'gulf'	golf	golf	golf	golf
	(b) *pont* 'bridge'	pɔn	pɔn	pɔnt	pɔnt
	camp 'field'	kam	kam	kamp	camp
	alt 'high.M'	al	al	alt	alt
	(c) *part* 'part'	part	paɾ	paɾt	paɾt
	gust 'taste'	gust	gus	gust	gust
	(d) *cranc* 'crab'	kɾaŋ(k)	kɾaŋ(k)	kɾaŋk	kɾanʲc
	carn 'meat'	karn	karn	kaɾn	caɾn

(2)	Derivatives of words in (1) revealing input form of stem	Central Catalan pronunciation	Morphological structure
	(a) *arquet* 'bow'	əɾˈkɛt	ˈark+ˈɛt
	serpeja 'snake.3SG.PR.IND'	səɾˈpɛʒə	ˈserp+ˈɛʒ+a
	golfada 'bad weather in a gulf'	gulˈfaðə	ˈgolf+ˈad+a
	(b) *pontet* 'bridge.DIM'	punˈtɛt	ˈpɔnt+ˈɛt
	campestre 'rural'	kəmˈpɛstɾə	ˈkamp+ˈɛstɾ
	alta 'high.F'	ˈaltə	ˈalt+a
	(c) *partir* 'divide'	pəɾˈti	ˈpart+ˈi+ɾ
	gustós 'tasty'	gusˈtos	ˈgust+ˈoz
	(d) *cranquet* 'crab.DIM'	kɾəŋˈkɛt	ˈkɾaŋk+ˈɛt
	carner 'meat-eating'	kəɾˈne	ˈkarn+ˈeɾ

The forms in (1a) show that 'final consonant deletion' does not affect heterorganic clusters. Those in (1b) show homorganic stops reduced after nasals and laterals in all styles in Catalonia, but retained in Valencian and Majorcan. In (1b) the members of the cluster share values of [continuant], i.e. the clusters are *partial geminates*, in that place and important aspects of manner are shared. However, the value of [continuant] for [l] is determined by adjacent segments; as

explained in §10.1.6, [l] is [−cont] only when adjacent to a homorganic (dentialveolar) stop. Therefore, with cluster reduction, the output [l] generally ceases to be in an environment where its [−cont] value is maintained. From this perspective, 'homorganic cluster simplification' is not quite the same process with lateral C_1 as with nasal C_1. In fact, while homorganic nasal cluster reduction is virtually categorical in Catalonia (see below), there is evidence that stop deletion in /lt/ clusters is to a degree lexically conditioned (see Wheeler (1981: 606) where, as far as I know, this aspect was first raised). Recasens (1993: 167) reports categorical deletion in *alt* 'high, tall', *malalt* 'ill', *molt* 'much, many', *salt* 'jump', and *dalt* 'above', which are described as 'words of high frequency of use'. In other /-lt/ words, stop deletion is described as more typical of colloquial pronunciation, but also as 'ever more general, so that the rule is not far from becoming categorical' (168).[1] Only the reduced variant is considered further here. In (1c) reduction is characteristic of less formal styles in Catalonia. In contrast with the (1b) types, the final stop in (1c) is preceded by a continuant. Variation here is probably indicative of change in progress, according to Recasens, and such change is more advanced in /rt/ than in /st/. It is worth quoting Recasens *in extenso* on this matter:

> By contrast with the clusters with nasal or lateral C_1, there are no words ending in *-st* or *-rt*/*-rd* in which an oral dental stop C_2 is obligatorily deleted. . . . Maintenance and deletion of C_2 may alternate in colloquial pronunciation, before a pause and before a following word-initial vowel. Certain frequently used lexemes delete the consonant often: (*-st*) *gust* ('taste'), *disgust* ('displeasure'), *trist* ('sad'), *agost* ('August'), *just* ('just'), *vist* ('seen'); (*-rt*, *-rd*) *tort* ('wrong'), *sort* ('luck'), *port* ('harbour'), *mort* ('dead'), *part* ('part'), *fort* ('strong'), *curt* ('short'), *julivert* ('parsley'), *hort* ('smallholding'), *surt* ('goes out'), *quart* ('fourth'), *cert* ('certain'), *card* ('thistle'), *perd* ('loses'), *llard* ('animal fat'), *tard* ('late'), *bord* ('wild'), *cort* ('court'), *verd* ('green'), *sord* ('deaf'), *nord* ('north'). In other more learned words maintenance of C_2 is favoured to a greater degree: (*-st*) *manifest* ('manifest'), *incest* ('incest'), *impost* ('tax'), *nefast* ('ill-fated'), *bust* ('bust'), *test* ('flower-pot', 'test'), *cast* ('chaste'), *cost* ('cost'); (*-rt*, *-rd*) *ert* ('stiff'), *confort* ('comfort'), *infart* ('infarct'), *absurd* ('absurd'), *sard* ('Sardinian'), *curd* ('Kurd'). Deletion is more frequent after the rhotic than after the fricative. In fact, the retention of the stop in several popular words ending in *-st* ([st]: *rast* ('string'), *llast* ('ballast'), *tast* ('taste'), *bast* ('coarse'), *abast* ('reach'), *llangost* ('locust'), *post* ('board'), *llest* ('clever'), *oest* ('west'), *prest* ('ready')) contrasts with its variable deletion in a good number of more or less popular words ending in *-rt*/*-rd* ([rt] ∼ [r]: *expert* ('expert'), *desert* ('desert'), *concert* ('concert'), *suport* ('support'), *passaport* ('passport'), *esport* ('sport'), *import* ('value'), *acord* ('agreement'), *record* ('memory')). The treatment of participles and [participial] adjectives is also indicative: deletion more often affects forms ending in *-rt* (*mort* ('dead'), *cobert* ('covered'),

[1] Recasens mentions as words with possible unreduced [lt#]: *contralt* 'contralto', *insult* 'insult', *adult* 'adult', *indult* 'pardon', *herald* 'herald', *cobalt* 'cobalt', *asfalt* 'asphalt', *esmalt* 'enamel'; and the participles *mòlt* 'ground', *absolt* 'absolved', and *resolt* 'resolved'. Other words where some sources report variable [lt] are *assalt* 'assault', *basalt* 'basalt', *esvelt* 'svelte', *ocult* 'concealed', *tumult* 'tumult'.

obert ['open, opened'], *ofert* ['offered'], *sofert* ['undergone'], *establert* ['established']) than forms ending in *-st* (*respost* 'replied'], *compost* ['composed'], *descompost* ['split up'], *previst* ['foreseen'], *post* ['laid']). (Recasens 1993: 168 (trans. MWW))

In the present analysis a retained stop is treated as representative of 'more formal style' and deletion as representative of 'less formal'.

The realization of velar (ŋk) coda clusters (1d) differs slightly from the other homorganic nasal clusters. Though deletion is widespread in this context in Catalonia, retention of the velar stop can also be heard in prepausal or prevocalic position (e.g. *en tinc cinc* [ən.ˈtiŋ.ˈsiŋk] ∼ [ən.ˈtiŋˈsiŋ] 'I have five of them'; *blanc i negre* [ˈblaŋ.ki.ˈnɛɣɾə] ∼ [ˈbla.ŋi.ˈnɛɣɾə] 'black and white'). According to Recasens (1993: 167) the factors contributing to variation here may be lexical, stylistic, and geographical (though he does not specify what the lexical and stylistic factors may be). However, the difference between the velar nasal clusters and the other homorganic ones no doubt reflects articulatory and perceptual factors. The slower motion of the dorsum compared with that of the lips and the front of the tongue may lead to the maintenance of oral occlusion after the opening of the velo-pharyngeal passage. And the greater duration and perceptibility of the stop burst of [k] would contribute to its retention being taken as more characteristic than in the case of the labial or dental stop (Recasens 1991b: 242). In the present account, absence of the velar stop after a homorganic nasal is taken as representative of both the more and the less formal styles of pronunciation in Catalonia. In (1d) *carn* is given as an example of a homorganic cluster (/rn/) that is not subject to deletion in Catalonia.[2]

The constraint responsible for coda cluster simplification is, not surprisingly, the markedness constraint *ComplexCoda, or, more precisely, a constraint of the *Coda family. I abbreviate the relevant constraint here as $*CC]_{\sigma}$ (3).

(3) $*CC]_{\sigma}$: A coda does not contain two segments.

For the (1b) examples it is clear that some markedness constraint against clusters outranks some faithfulness constraint that penalizes reduction in the number of segments, whether Max, the correspondence constraint which penalizes deletion, or Uniformity (§6.5 (45)), which penalizes fusion. However, the (1a) examples and (1d) *carn* [karn] 'meat', and for the more formal style the (1c) examples, show that some faithfulness constraints outrank $*CC]_{\sigma}$. For the Catalonia more formal variety it will be shown that the relevant faithfulness constraints are IdentI-OPlace (IdentPA) (4), IdentI-OManner (5), and IdentI-OSibilant (6). The approach to featural identity within correspondence theory taken here derives from Pater (1999, 2001). One thing that follows from Pater's approach is that successful homorganic cluster reduction may be interpreted as fusion rather than as deletion in the strict sense (involving a Max violation).

[2] Other such clusters are /ls/ (*fals* 'false'), /ns/ (*tens* 'tense'), /nz/ (*fons* 'bottom'), /rs/ (*terç* 'regiment'), /mf/ (*triomf* 'triumph'), /nʃ/ (*ponx* 'punch'), /lʃ/ *Elx* (toponym), and /tʃ/ (if taken to be a cluster, e.g. *despatx* 'office').

(4) IDENTI-OPlace (IDENTPA): Any correspondent of an input segment specified for place of articulation must have the same specification.

(5) IDENTI-OManner (IDENTMan): Any correspondent of an input segment specified for manner of articulation must have the same specification. ('Manner' denotes one or more of [sonorant] [continuant], Sibilant, Nasal, Lateral, and Rhotic; each of these forms the content of a more specific manner faithfulness constraint, as in (6).)

(6) IDENTI-OSibilant (IDENTSib): Any correspondent of an input segment specified for Sibilant must have the same specification.

Cluster reduction evidently affects the exterior element of the cluster. This follows most obviously from the activity of the CONTIGUITY constraint (7), though other relatively high faithfulness constraints such as IDENTNasal, IDENTRhotic also preserve interior elements. The activity of CONTIGUITY is assumed in the tableaux that follow in (9,10).

(7) CONTIGUITY (CONTIG): The portion of S_1 standing in correspondence forms a contiguous string (simplified from McCarthy & Prince 1995: 123).

As mentioned above, the issue with homorganic nasal–stop clusters and with homorganic lateral–stop clusters is slightly different. From the perspective of the input, both types of cluster consist of partial geminates (shared place, and shared [−cont]). Thus in a rule-based theory the clusters could be dealt with in exactly the same manner: the structural description is the same, and so is the structural change (whether fusion or deletion). But Optimality evaluates outputs, and it is here that nasal and lateral cluster reduction differs; the nasal output of the nasal clusters remains [−cont], but the output of the lateral cluster is [+cont], unless, fortuitously, the lateral is followed on the surface by a denti-alveolar stop. Thus, for *pont* 'bridge', both [ˈpɔn$_{12}$] and [ˈpɔn$_1$t$_2$] corresponding to /ˈpɔn$_1$t$_2$/ observe IDENT[cont], and the former is preferred by the *COMPLEXCODA constraint *CC]$_\sigma$. The same is not the case for *alt* realized as [ˈal]. Here, absence of the stop, by fusion, leads to a violation of IDENT[cont]. In this style the retention of [t] in *part* 'part' and *gust* 'taste' shows that IDENT[cont] outranks *CC]$_\sigma$, hence reduction in *alt* must be driven by some other principle. I suggest that what is evident in the Catalonia more formal style is a constraint specifically against a partial geminate (or, indeed, any kind of geminate) in a coda. The motivation would be primarily perceptual. The members of a cluster, sharing place and continuancy, are not sufficiently different to be adequately distinguishable, so the attempt to realize the cluster is abandoned. I label this constraint *GEMINATECODA (8).

(8) *GEMINATECODA (* $C_iC_i]_\sigma$): Adjacent consonants in a coda do not share values of both place and manner (where 'manner' denotes one or more of [cont], Sibilant, Nasal, Lateral, Rhotic).

Tableaux (9) and (10) show the relation between the constraints mentioned so far in the more formal style of Catalonia, with homorganic clusters in (9) and heterorganic clusters in (10). The examples in (9) are *camp* 'field', *alt* 'high', *part* 'part', *carn* 'meat', and *fals* 'false'. In (10) they are *serp* 'snake' and *golf* 'gulf'.

(9)

	Catalonia more formal		*$C_iC_i]_\sigma$	IDSib	IDNas	IDLat	IDRhot	IDPA	MAX	ID[cont]	IDMan	*$CC]_\sigma$	UNIF
(a)	*cam_1p_2*	kam_1p_2	!*									*	
		kam_1							*!				
		kan_{12}						*!					
		kap_{12}			*!						*		*
	☞	kam_{12}									*		*
(b)	*al_1t_2*	al_1t_2	*!									*	
		al_1							*!				
		at_{12}				*!					*		*
	☞	al_{12}								*	*		*
(c)	*par_1t_2*	☞ par_1t_2										*	
		par_1							*!				
		pat_{12}					*!			*	*		*
		par_{12}								*!	*		*
(d)	*car_1n_2*	☞ kar_1n_2										*	
		kar_1							*!				
		kan_{12}					*!			*	*		*
		kar_{12}			*!					*	*		*
(e)	*fal_1s_2*	☞ fal_1s_2										*	
		fal_1							*!				
		fas_{12}				*!					*		*
		fal_{12}		*!							*		*

(Catalonia more formal) *$C_iC_i]_\sigma$, IDENTSib, IDENTNas, IDENTLat, IDENTRhot, IDENTPA, MAX » IDENT[cont], IDENTManner » *$CC]_\sigma$, UNIFORMITY

The faithful candidates of (9a) and (9b) (line 1) have fatal partial geminates in a coda, so an alternative with cluster reduction is preferred, faithful to Nasal or Lateral. When partial geminates are not involved, (9c) and (9d) show that MAX and/or IDENT[cont], or IDENTRhotic, outrank $*CC]_\sigma$. Neither deletion nor fusion that does not respect [continuant] is an acceptable improvement on a Complex Coda violation. Example (9e), where IDENT[cont] is not involved, shows specifically that MAX outranks $*CC]_\sigma$, since no other constraint distinguishes the first two candidates. In the style of speech that allows final [lt], the *GEMINATECODA constraint lies below MAX and IDENT[cont], and de facto ceases to be distinguishable from $*CC]_\sigma$. In this variety, I suggest words with [-l]/[-l.t-] alternation (*molt* and so on) display allomorphy that is both lexically and phonologically conditioned (as happens with other cases of 'level 1' deletion (§10.2)).

Tableau (10) shows that faithfulness to place (IDENTPA) discounts the fusion candidates for (1a) examples (since place is not retained in them), while MAX weeds out the deletion candidates. (Candidates violating CONTIGUITY, i.e. $*[sep_2]$, $*[gof_2]$ would fail by the same constraints, while fusion candidates $*[sep_{12}]$, $*[gof_{12}]$ would violate IDENTRhotic and IDENTLateral respectively.)

(10)

	Catalonia more formal	$*C_iC_i]_\sigma$	IDSib	IDPA	MAX	ID[cont]	IDMan	$*CC]_\sigma$	UNIF
(a)	*ser*$_1$*p*$_2$ ☞ ser$_1$p$_2$							*	
	ser$_1$				*!				
	ser$_{12}$			*!		*			*
(b)	*gol*$_1$*f*$_2$ ☞ gol$_1$f$_2$							*	
	gol$_1$				*!				
	gol$_{12}$			*!					*

There is evidence, by the way, that $*CC]_\sigma$ in fact ranks above UNIFORMITY. This evidence comes from the affrication of denti-alveolar+sibilant clusters, provided that an affricate is interpreted as a contour segment with [[−cont][+cont]]. So for *tots* /tot$_1$+z$_2$/ 'all.PL' the output [tot$\widehat{\text{s}}_{12}$] with an affricate wins over [tot$_1$s$_2$], as the latter violates $*CC]_\sigma$ while the former only violates the lower ranking UNIFORMITY. Henceforth I assume $*CC]_\sigma$ » UNIFORMITY.

The data in (1) shows that the Complex Coda constraint plays a much less important role in Valencian and Majorcan, where words with -CC# input are pronounced faithfully. For these varieties IDENTManner and MAX outrank $*CC]_\sigma$, as shown in (11).

(11)

	Valencia and Majorca	ID[cont]	IDManner	MAX	*CC]$_\sigma$	UNIF
(a)	*pon*$_1$*t*$_2$ ☞ pɔn$_1$t$_2$				*	
	pɔn$_1$			*!		
	pɔt$_{12}$		*!			
	pɔn$_{12}$		*!			*
(b)	*par*$_1$*t*$_2$ ☞ paɾ$_1$t$_2$				*	
	paɾ$_1$			*!		
	pat$_{12}$		*!			
	paɾ$_{12}$	*!	*			*

(Valencia/Majorca) IDENTSib, IDENT[cont], IDENTManner, MAX » *CC]$_\sigma$ » UNIFORMITY

Example (11a) shows the role of the IDENTManner constraint; the fusion candidate [pɔn$_{12}$] respects IDENT[cont]—the value for [continuant] of each corresponding segment /n/ and /t/ is preserved—but is still ungrammatical in these varieties. (No doubt IDENTNas, IDENTLat, IDENTRhot rank at least as high as IDENT[cont] in Majorca and Valencia.)

7.3 CLUSTER REDUCTION IN STEMS WITH A CONSONANTAL SUFFIX

Additional complexity is introduced to codas through consonantal inflectional suffixes, of which the most important is plural /+z/. (Stems with second-person singular suffix /+z/ are realized in exactly the same way and are not separately illustrated here. The matter of consonantal enclitics is taken up in Chapter 11.) The types are illustrated in (12) with their realization in the three regional dialects that are being compared in this section. Observe that in the case of final /-sC/ clusters, as in *gust* 'taste', an allomorph of the (masculine) plural suffix is available: /+oz/. This allomorph, which avoids many coda complexity problems, is universally preferred in Catalonia, and is used in some parts of Valencia, but not in Majorca.

(12)

Plural forms of stems in (1)	Catalonia	Valencia	Majorca
(a) *arcs* 'arches'	arks	aɾks	aɾs
serps 'snakes'	serps	seɾps	seɾs
golfs 'gulfs'	golfs	golfs	gols
(b) *ponts* 'bridges'	pɔns	pɔns	pɔns
camps 'fields'	kams	kams	cans
alts 'high.M.PL'	als	als	als

Plural forms of stems in (1)	Catalonia	Valencia	Majorca
(c) *parts* 'parts'	pars	parts͡	pars
gusts ∼ *gustos* 'tastes'	gustus	gusts͡	guts͡
(d) *crancs* 'crabs'	kraŋs	kraŋs	krajns
carns 'meats'	karns	karns	cans

Adding /+z/ gives rise to structures which are phonologically marked, from several distinct points of view. In the first place, adding a consonant to a complex coda will result in one which is yet more complex: $CCC]_{\sigma}$. In Majorca the constraint against tri-consonantal codas is very highly ranked, with the reduction results seen in (12). Secondly, adding an alveolar /z/ may give rise to heterorganic clusters (as in *serps* /serp+z/ 'snakes') against which several constraints may conspire, especially in Majorca, as seen in Chapter 6; thus in Majorca *camps* 'fields' is realized with place assimilation as well as cluster reduction: [cans]. Thirdly, a fricative in a coda after another obstruent gives rise to a structure that violates the sonority sequencing principle (SonSeq). Thus, for example, the realization of *alts* /alt+z/ 'high.PL' as [alts] involves a SonSeq violation; sonority falls from [a] via [l] to [t], then rises again to [s]. It will be seen that SonSeq plays at most a minor role in cluster reduction. And as inflection here involves addition to a base, the matter of Paradigm Uniformity arises. That is, we may expect to see pressure to realize a stem in the same way before /+z/ as it is realized in its base form. Lastly, the presence of an inflectional morpheme may bring to view a MaxMorpheme (MaxM) constraint penalizing the deletion of semantically relevant material.

In the context illustrated in (12) there is no difference in Catalonia between a more formal and a less formal style. However, the forms in (12) do not relate to the base forms in the two styles in (1) in the same way. While the forms are the same, the constraint ranking is different. In the singular (more formal style) reduction is found with homorganic partial geminate clusters (*pont* [pɔn], *alt* [al]) but not with other homorganic clusters (*part* [part], *carn* [karn]). In the plural forms (before /+z/), however, *parts* shows reduction ([pars]) analogously to *ponts* [pɔns] and *alts* [als], while *carns* [karns] remains unreduced. One reason might be that *[parts], unlike [karns], displays a Sonority Sequence violation, as mentioned above. However, it will be seen later that the same pattern of reduction and non-reduction is found when *part* and *carn* stand before consonant-initial words (*part baixa* 'lower part' [ˈpaɾ.ˈβaʃə] but *carn freda* 'cold meat' [karn.ˈfrɛðə]) and in this context there is no SonSeq violation. What drives cluster resolution here, rather, is the higher position of Complex Coda constraints such as $*CCC]_{\sigma}$. In fact, in many cases it seems that the crucial Complex Coda constraint is not simply one against the presence of three consonants in a coda (though that would justify, e.g. for *camps* 'fields', the preference for [kams] over [kamps]). Rather, the constraint is against the complexity consisting of two

changes of manner of articulation in a coda, so, for *alts* 'high.M.PL', [als] is even better than [al͡ts]; each has a CC]$_\sigma$ coda, but the latter's has a [−cont] [+ cont][−cont] sequence. I label this the THREE-MANNER CODA constraint as in (13).

(13) *THREE-MANNER CODA (*3MAN]$_\sigma$): There is no more than one point where change of Manner of Articulation occurs in a coda (where Manner means Rhotic, Nasal, Sibilant, Lateral, [±continuant] or [±sonorant]).[3]

The tableau in (14) compares various candidates for *arcs* 'arches' in the constraint hierarchy relevant to the more formal style in Catalonia.

(14)

	Catalonia more formal ar_1k_2+z sg. [ark]	IDSib	IDPA	MAX	*3MAN]$_\sigma$	*CCC]$_\sigma$	SONSEQ	IDMan	MAXM	PU	ID[cont]
(a)	☞ arks				*	*	*				
(b)	ar_1s			*!						*	
(c)	ar_{12}s		*!					*		*	*
(d)	ark_1			*!					*		
(e)	ark_{12}	*!	*					*	*		*

(Catalonia more formal) IDENTSib, IDENTPA, MAX » *3MAN]$_\sigma$, *CCC]$_\sigma$, SONSEQ, IDENTManner, MAXMorph, PARADIGMUNIFORMITY, IDENT[cont]

In (14) PARADIGMUNIFORMITY penalizes those candidates which do not contain [ark], the form of the base of the inflected form. The winner, (14a), violates both SONSEQ and *THREE-MANNER CODA as well as the tri-consonantal version (*CCC]$_\sigma$) of the Complex Coda constraint. The role of MaxM here might suggest that final /s/ or /z/ in clusters could be deleted if it did not express a morpheme. There are a few such words, such as *linx* [liŋs] 'lynx', but in such words the presibilant sequence is homorganic, hence /Nk/ can be realized [ŋ] without further reduction taking place. (The plural of *linx* is *linxs* [liŋs], i.e. homophonous with the singular. As elsewhere, the constraint *GEMSib disfavours a realization with a geminate: *[liŋss].) Henceforth I ignore MAXM, whose effects are covered by IDENTSib, a constraint active for other reasons.

In (15) are displayed candidates for *crancs* 'crabs', *parts* 'parts', and *carns* 'meats' in the same more formal variety of Catalonia. IDENTSib is left out to simplify the exposition. On the other hand, IDENTNas is introduced. Although unviolated in any winning form, it is relevant to the candidates for *carns*.

[3] This constraint is offered provisionally. Further investigation is needed. Conceptually the parallel is with other types of articulatory complexity in consonant clusters, such as heterorganic place, or voicing contours.

(15)

	kraN$_1$k$_2$+z sg. [kraŋ]	IDNAS	IDPA	MAX	*3MAN]$_\sigma$	*CCC]$_\sigma$	SONSEQ	IDMan	PU	UNIF
(a)	kraŋks				*!	*	*		*	
	kraŋ$_1$s			*!						
	kran$_1$s		*!	*					*	
	☞ kraŋ$_{12}$s									*
(b)	part+z sg. [part]									
	parts				*!	*	*			
	parts͡				*!		*			*
	par$_1$s			*!					*	
	☞ par$_{12}$s							*	*	*
(c)	karn+z sg. [karn]									
	☞ karns				*	*				
	kar$_1$s			*!					*	
	kar$_{12}$s	*!						*	*	*

(Catalonia more formal) IDENTNAS, IDENTPA, MAX » *3MAN]$_\sigma$, *CCC]$_\sigma$, SONSEQ » IDENTManner, PU, UNIFORMITY

In (15a) and (15b) formally the fused candidates win over the deletion ones, due to dominance of MAX, but since these are not phonetically distinct it would not matter here if MAX and UNIFORMITY switched places. In (15a) I have registered an IDENTPA violation against [kran$_1$s], which is perhaps not formally correct, as place is not contrastive in a pre-velar nasal.[4] This candidate falls to the winner on other grounds anyway. Observe that the winner in (15b) violates PARADIGM UNIFORMITY—the singular is [part] in the more formal style. In the less formal style the outcome is the same, though in this variety the plural form candidates [par$_1$s]/[par$_{12}$s] incur no PARADIGM UNIFORMITY violations—the singular is [par].

The Valencian suffixed forms (as described by Jiménez 1999: 213) differ from those in Catalonia only in that [t] is retained in *parts* [parts͡] and *gusts* [gusts͡] 'tastes'.[5] It is not obvious why, given singulars *alt* [alt] 'high' and *part* [part]

[4] As Clàudia Pons has pointed out to me (p.c.), this difficulty is avoided if the relevant IDENTPA constraint is understood to be an output-output correspondence constraint; the required faithfulness is to the PA of the singular base form *cranc* [ˈkraŋ].

[5] Other Valencian varieties have [pars], as in Catalonia; or stops may be retained in all the clusters of (12) giving [pɔnts͡], [alts͡], etc. (Recasens 1991b: 346–350). In the latter varieties PARADIGM UNIFORMITY dominates *CCC]$_\sigma$ or *3MAN]$_\sigma$ without further ado.

'part', the Valencian plurals should be [als] but [paɾt͡s] respectively. Jiménez's brief discussion of *alts* (1999: 222) puts *alts* alongside *ponts* [pɔns] as cases in which both place and 'manner' (i.e. continuancy) are preserved in the reduced cluster. This makes sense only if he regards [l], in this context ([ls]) at least, as being a non-continuant, though generally [l] is regarded as [−cont] only when adjacent to a dental stop [t] or [d] (see §10.1.6). Thus, presumably, [l] behaves like a non-continuant when adjacent to any denti-alveolar consonant (or even any coronal), given that tongue-front contact for the lateral interrupts the airflow for any adjacent tongue-front consonant. In the interpretation of Valencian I offer here I assume this to be the case. Recasens (1991b: 350), however, describes [lts] as a major Valencian variant (alongside [ls]). For Recasens, the /lts/ cluster in Valencian differs from the homorganic nasal clusters (/nts/, /mps/, at least); in the latter 'the presence of an oral stop is secondary and not very frequent'. It looks as though for Recasens, then, realization in Valencian of /lts/ is parallel to that of /rts/, while realization of /nts/ and /mps/ (homorganic stop/partial geminate clusters) is different. Further investigation is required. Tableau (16) illustrates the evaluation of candidates for *arcs* 'arches', *alts* 'high.M.PL', and *parts* 'parts' in Valencian.

In Valencian *parts* is treated like *carns* 'meats' but differently from *alts*. Thus in this variety IDENT[cont] (which prefers [paɾt] to [paɾ] and [paɾt͡s] to [pars]) is up alongside IDENTNas and IDENTPA. Here too SONSEQ[6] OR $*3\text{MAN}]_\sigma$ is active (16b) in preferring [als] to [alt͡s], despite the latter being a better candidate from the perspective of PARADIGM UNIFORMITY. PARADIGM UNIFORMITY is located lower than SONSEQ and the Complex Coda constraint $*3\text{MAN}]_\sigma$ in Valencian, allowing reduced candidates for partial geminate clusters (*ponts* [pɔns] 'bridges') to surface in the plural despite their being preserved in the singular (*pont* [pɔnt]).

(16)

(a)

Valencia $aɾ_1k_2+z_3$ sg. [aɾk]	IDNas	IDRhot	IDSib	IDPA	ID[cont]	MAX	SONSEQ	$*3\text{MAN}]_\sigma$	PU	IDMan	UNIF
☞ aɾks							*	*			
$aɾ_1s$						*!			*		
$aɾ_{12}s$				*!	*				*	*	*
$aɾk_2$						*!					
$aɾk_{23}$			*!		*					*	

[6] Pending further investigation of the place of affricates in the sonority hierarchy, here SONSEQ is given a restrictive interpretation, such that coda affricates are penalized in the same way as stop + fricative clusters.

	al$_1$t$_2$+z sg. [alt]	IDNas	IDRhot	IDSib	IDPA	ID[cont]	MAX	SONSEQ	*3MAN]$_\sigma$	PU	IDMan	UNIF
(b)	al$\widehat{\text{ts}}$							*!	*			
	al$_1$s						*!			*		
	☞ al$_{12}$s									*	*	*
(c)	par$_1$t$_2$+z sg. [part]											
	☞ par$\widehat{\text{ts}}$							*	*			
	pa$\widehat{\text{ts}}_{123}$		*!					*		*	*	
	par$_1$s						*!			*		
	par$_{12}$s					*!				*	*	*

(Valencia) IDENTNas, IDENTRhotic, IDENTSib, IDENTPA, IDENT[cont], MAX » *3MAN]$_\sigma$ » PU, IDENTMan, UNIFORMITY

Two aspects are notably characteristic of the Majorcan resolution of tri-consonantal codas in (12). The first is substantial reduction, with elimination of the stop consonant. The second is major place assimilation to the suffix /+z/ so no heterorganic clusters remain. Another peculiarity is 'compensatory diphthongization'—the resolution of nasal+dorsal+sibilant /Nc+z/ as [jns]. Note that, exceptionally, this resolution results in a tri-consonantal coda (CCC]$_\sigma$, but not a three-manner coda, assuming manner is only relevant to [+cons] segments) such as is elsewhere strictly avoided in suffixed forms. Lastly, the Majorcan resolution of /ɾn+z/ clusters as in *carns* [cans] 'meats' involves the disappearance not of the middle consonant, as elsewhere, but of the first, i.e. the rhotic, despite the surface CONTIGUITY violation. What drives Majorcan cluster resolution is the higher position of Complex Coda constraints, in particular the *THREE-MANNER CODA constraint, against the complexity consisting of two changes of manner of articulation in a coda. Though it might appear that what is crucial here is the SONSEQ violation in [al$\widehat{\text{ts}}$], in other cases an affricate (violating SONSEQ) wins over a fricative, namely in *gusts* 'tastes', where [gu$\widehat{\text{ts}}$] is preferred to [gus].

Assimilation, as in *camps* [cans], reflects the lower position of IDENTPA in this variety. To show how this works a wider range of candidates needs to be considered than in the tableaux given so far in this section. Tableau (17) considers a range of candidates for *arcs* 'arches' in Majorcan.

(17)

	Majorca $ar_1c_2+z_3$ sg. [arc]	*3MAN]$_\sigma$	MAX	IDSIB	SONSEQ	*CCC]$_\sigma$	IDMAN	*CC]$_\sigma$	IDPA	PU	UNIF	INTEG
(a)	arcs	*!			*	*		**				
(b)	ar_1s		*!					*		*		
(c)	☞ $ar_{12}s$						*	*	*	*	*	
(d)	arc_2		*!					*				
(e)	arc_{23}			*!			*	*	*		*	
(f)	art_{23}			*!	*			*	*	*	*	
(g)	☞ ars_{23}						*	*	*	*	*	
(h)	$ac_{12}s$				*		*!	*	*	*	*	
(i)	$ar\widehat{ts}_{23}$	*!			*			*	*	*	*	
(j)	ar_{123}			*!			*		*	*	**	

(Majorca) *3MAN]$_\sigma$, MAX, IDENTSIB » SONSEQ, *CCC]$_\sigma$, IDENTMan » *CC]$_\sigma$ » IDENTPA, PU, UNIFORMITY, INTEGRITY

Both (17a) and (17i) violate *THREE-MANNER CODA, though only the former violates *CCC]$_\sigma$. Candidates (17e) and (17f) attempt to repair the *THREE-MANNER CODA violation by fusing the last two consonants as a stop, but these fall to IDENTSib (or MAXM). Strict deletion candidates (17b) and (17d) are weeded out by MAX. (17h) fatally violates SONSEQ (also CONTIGUITY). There are two winners, (17c) and (17g), which do not differ in pronunciation; each shows fusion—in (17c) of C_1 and C_2, in (17g) of C_2 and C_3.

Tableau (18) evaluates possible pronunciations of *crancs* 'crabs' in Majorcan, where the winner is [krajns] with 'compensatory diphthongization'.

(18)

	Majorca $kraN_1c_2+z_3$ sg. [$kran^jc$]	*3MAN]$_\sigma$	IDNAS	IDPAL	*{ʃKPT}$_{[-ct]}${ʃKPT}	MAX	IDSIB	SONSEQ	*CCC]$_\sigma$	IDMAN	IDPA	PU	UNIF	INTEG
(a)	$kran^jcs$	*!			*			*	*					
(b)	$kran^j{}_{12}s$				*!					*		*	*	
(c)	$kran^j{}_1s$				*!	*						*		
(d)	$kran_{12}s$			*!						*	*	*	*	
(e)	$kran_1s$					*!					*	*		
(f)	$kran_1\widehat{ts}_{23}$	*!		*				*			*	*	*	
(g)	☞ $kraj_{12}n_{12}s$									*		*	**	**
(h)	$kraj_1n_2s$		*!	*						*	*	*		

(Majorca) *3MAN]σ, IDENTNas, IDENTPal, *{ʃKPT}$_{[-ct]}${ʃKPT}, MAX, IDENTSib » SONSEQ, *CCC]$_\sigma$, IDENTMan » *CC]$_\sigma$ » IDENTPA, PU, UNIFORMITY, INTEGRITY

As in (17), the faithful candidate (18a) is debarred for a *THREE-MANNER CODA violation. All the other candidates, of course, incur a PARADIGM UNIFORMITY violation since the form of the singular is not copied in the plural. Candidate (18h) (which is pronounced the same as the winner) has no fusion or splitting, but its $[j_1n_2]$ is not faithful to input $[n^j{}_1c_2]$ in nasality or place. As before, the strict deletion candidates (18c) and (18e) are excluded by MAX. (18b), which shows fusion of sonorant and stop as a sonorant ($[n^j]$) identical with the one found in the singular form, matches in structure the winning candidate in (17) — (17c) — but in this case (18b) falls to $*\{\text{ʃKPT}\}_{[-\text{ct}]}\{\text{ʃKPT}\}$, the strong Majorcan heterorganic cluster constraint discussed in Chapter 6. The winner (18g) contains two multiply correspondent segments, hence two UNIFORMITY violations (two cases of fusion: $C_1C_2 \rightarrow C_{12}$), and two INTEGRITY violations (two cases of splitting: $C_i \rightarrow C_iC_i$).

The evaluation for Majorcan *gusts* 'tastes' (19) is particularly complex, as there are many conceivable alternatives to faithful implementation, especially when possible fusion candidates are considered.

(19)

	Majorca $gus_1t_2+z_3$ sg. [gust]	$*3\text{MAN}]_\sigma$	$*\{\text{ʃKPT}\}_{[-\text{ct}]}\{\text{ʃKPT}\}$	MAX	IDSib	$*\text{CCC}]_\sigma$	IDMan	SONSEQ	$*\text{CC}]_\sigma$	IDPA	PU	UNIF	INTEG
(a)	$gus_1t_2s_3$	*!				*		*	**				
(b)	$gus_1\widehat{ts}_{23}$	*!						*	*			*	
(c)	$gu\widehat{ts}_{23}$			*!				*			*	*	
(d)	☞ $gu\widehat{ts}_{123}$							*			*	**	
(e)	$gut_{12}s_3$				*!		*	*	*		*	*	
(f)	gus_1t_2			*!					*				
(g)	gus_1t_{23}				*!		*		*			*	
(h)	gus_{123}						*!				*	**	
(i)	gut_{123}				*!		*				*	**	
(j)	gus_1s_3			*!					*		*		
(k)	$gus_{12}s_3$						*!		*		*	*	

(Majorca) $*3\text{MAN}]_\sigma$, $*\{\text{ʃKPT}\}_{[-\text{ct}]}\{\text{ʃKPT}\}$, MAX, IDENTSib » $*\text{CCC}]_\sigma$, IDENTMan » SONSEQ » $*\text{CC}]_\sigma$, » IDENTPA, PU, UNIFORMITY, INTEGRITY

The candidates without a final sibilant in (19)—(19f), (19g), (19i)—fall either to MAX or to IDENTSib. Candidate (19e) violates IDENTSib in its C_1 correspondent. Those without a plosive or affricate in the coda—(19h), (19j), (19k)—fall to

IDENTManner, which is seen to rank above SONSEQ as [t͡s.] in the winner incurs a SONSEQ violation. The winner, (19d), has multiple fusions (hence two asterisks under UNIFORMITY). The story for *carns* 'meats' (20) is not greatly different, though noteworthy here is IDENTNasal » IDENTManner, whereby (20c) (despite violating in particular IDENTRhotic) wins over (20b) with an IDENTNasal violation. The combined constraint hierarchy for (17–20) is given in (21).

(20)

	Majorca car$_1$n$_2$+z$_3$ sg. [carn]	*3MAN]$_\sigma$	IDNas	IDPal	*{ʃKPT}$_{[-cd]}${ʃKPT}	MAX	IDSib	*CCC]$_\sigma$	IDMan	SONSEQ	IDPA	PU	UNIF	INTEG
(a)	car$_1$n$_2$s$_3$	*!						*						
(b)	car$_{12}$s$_3$		*!						*			*	*	
(c)	☞ can$_{12}$s$_3$								*			*	*	
(d)	carn$_{23}$						*!		*					
(e)	carn$_2$					*!								
(f)	car$_1$s$_3$					*!						*		
(g)	can$_2$s$_3$					*!						*		

(21) (Majorca) *3MAN]$_\sigma$, IDENTNas, IDENTPal, *{ʃKPT}$_{[-ct]}${ʃKPT}, MAX, IDENTSib » *CCC]$_\sigma$, IDENTMan » SONSEQ » *CC]$_\sigma$ » IDENTPA, PU, UNIFORMITY, INTEGRITY

7.4 CLUSTER REDUCTION IN UNINFLECTED STEMS FOLLOWED BY VOWEL-INITIAL WORDS

In (22) are displayed the realizations of the words we have been considering in their base forms followed by vowel-initial words.

(22)

Words in (1) followed by vowel	Catalonia more formal	Catalonia less formal	Valencia	Majorca
(a) *arc ample* 'wide arch'	ˈar.ˈkam.plə	ˈar.ˈkam.plə	ˈar.ˈkam.ple	ˈar.ˈcam.plə
serp àgil 'quick snake'	ˈser.ˈpa.ʒil	ˈser.ˈpa.ʒil	ˈser.ˈpa.d͡ʒil	ˈser.ˈpa.ʒil
golf ample 'wide gulf'	ˈgol.ˈfam.plə	ˈgol.ˈfam.plə	ˈgol.ˈfam.ple	ˈgol.ˈfam.plə

(22)

Words in (1) followed by vowel	Catalonia more formal	Catalonia less formal	Valencia	Majorca
(b) *pont ample* 'wide bridge'	ˈpɔ.ˈnam.plə	ˈpɔ.ˈnam.plə	ˈpɔn.ˈtam.ple	ˈpɔn.ˈtam.plə
camp ampli 'broad field'	ˈka.ˈmam.pli	ˈka.ˈmam.pli	ˈkam.ˈpam.pli	ˈcam.ˈpam.pli
alt i gros 'tall.M and large'	ˈa.li.ˈɣrɔs	ˈa.li.ˈɣrɔs	ˈal.ti.ˈɣrɔs	ˈal.ti.ˈɣrɔs
(c) *part igual* 'equal part'	ˈpar.ti.ˈɣwal	ˈpa.ɾi.ˈɣwal	ˈpar.ti.ˈɣwal	ˈpar.ti.ˈɣwal
gust agre 'sour taste'	ˈgus.ˈta.ɣrə	ˈgu.ˈza.ɣrə	ˈgus.ˈta.ɣre	ˈgus.ˈta.ɣrə
(d) *cranc artificial* 'artificial crab'[7]	ˈkra.ŋər. ti.fi.ˈsjal	ˈkra.ŋər. ti.fi.ˈsjal	ˈkraŋ.kar. ti.fi.ˈsjal	ˈkranʲ.cər. ti.fi.si.ˈal
carn assecada 'dried meat'	ˈkar.nə.sə. ˈka.ðə	ˈkar.nə.sə. ˈka.ðə	ˈkar.na.se. ˈka.ða	ˈcar.nə.sə. ˈca.ðə

In this phrasal context the correspondence constraint IDENTPWd-PPhr (IDENTW-P) comes into consideration—it is the constraint penalizing deviation from the citation forms of words. What is most evident in the (22) forms (all varieties) is resyllabification, whereby word-final consonants become onsets. That is, ONSET ranks above UNIFORMσ ((15) in §3.1.1), the correspondence constraint making the syllable alignment of segments in phrases match that in their component words. In Catalonia, IDENTPWd-PPhr (IDENTW-P) outranks IDENTI-OManner, so we have prevocalic *pont* [pɔn], *alt* [al], just as in the singular, where the reduced forms respond to *GEMINATECODA (8) » IDENTI-OManner. In the other varieties, in relation to the data in (22), the ranking of IDENTW-P and IDENTManner is not crucial. The segments are uttered faithfully; only the syllabification differs.

In Catalonia there are some limited lexical exceptions to homorganic final cluster reduction in prevocalic contexts (Wheeler 1975/9: 281). The numeral *vint* is pronounced [bint] before the conjunction *i* 'and' in the numbers from 'twenty-one' to 'twenty-nine': *vint-i-cinc* [ˌbin.ti.ˈsiŋ(k)] 'twenty-five'. Though the structure is transparent and compositional, it is plausible that these numerals were lexicalized before the historical process of cluster reduction spread to prevocalic environments. (Such an effect is also visible in a few compounds like *cent-en-grana* [ˌsen.təŋ.ˈgranə] 'rupture-wort', *frontample* [ˌfronˈtamplə] 'having a broad forehead' (Bonet & Lloret 1998: 111).) *Cent* '100' also retains its [t] in some phrases; the sources mention *cent anys* '100 years' and *cent homes* '100 men'. *Sant* 'Saint' retains its [t] before the names of some vowel-initial saints: *Sant Andreu, Sant Antoni, Sant Esteve, Sant Isidre*, but not before all, and not before other words, for example *Sant Imperi* [ˌsanimˈpɛɾi] 'Holy Roman Empire'. The interrogatives *on* /ont/ 'where' and *quant* 'how much' may retain their [t] before frequently collocated verb forms, such as *on estava?*

[7] As mentioned in §7.2, I ignore here the Catalonia variant with velar stop retention [ˈkraŋ.kər.ti.fi.ˈsjal].

[ˌon.təs.ˈtaβə] 'where was it?' *quant és?* [ˌkwan.ˈtes] 'how much is it?' These contexts are the last relics of the last context (prevocalic, frequent collocations—on which see Bybee 2001: ch. 7) to which the lexical diffusion of homorganic nasal cluster reduction spread. Synchronically they may display allomorphy, for example /sant/ ∼ /sante/, in which the vowel-final allomorph would be expected to lose its final vowel (but not its last consonant) before a vowel-initial word.

7.5 CLUSTER REDUCTION IN UNINFLECTED STEMS FOLLOWED BY CONSONANT-INITIAL WORDS

I now consider final clusters in uninflected bases followed by consonant-initial words ($C_1C_2\#C_3$), illustrated in (23). There are natural similarities with the C_1C_2+C inflectional context, but the overall outcome is truly parallel only in the more formal pronunciation of Catalonia. The other varieties show more substantial reduction, with obstruent stop place and manner suppressed.

(23)	Words in (1) followed by consonant	Catalonia more formal	Catalonia less formal	Valencia	Majorca
(a)	*arc baix* 'low arch'	ˈarg.ˈbaʃ	ˈaɾ.ˈβaʃ	ˈaɾ.ˈβajʃ	ˈaɾ.ˈβaʃ
	serp gran 'large snake'	ˈserb.ˈɡran	ˈseɾ.ˈɣran	ˈseɾ.ˈɣran	ˈseɾ.ˈɣran
	golf gran 'large gulf'	ˈɡolv.ˈɣran	ˈɡol.ˈɣran	ˈɡol.ˈɣran	ˈɡol.ˈɣran
(b)	*pont baix* 'low bridge'	ˈpɔm.ˈbaʃ	ˈpɔm.ˈbaʃ	ˈpɔm.ˈbajʃ	ˈpɔm.ˈbaʃ
	camp nou 'new field'	ˈkam.ˈnɔw	ˈkam.ˈnɔw	ˈkam.ˈnɔw	ˈcan.ˈnɔw
	alt càrrec 'high position'	ˈal.ˈka.rək	ˈal.ˈka.rək	ˈal.ˈka.rek	ˈal.ˈca.rec
(c)	*part baixa* 'lower part'	ˈpaɾ.ˈβaʃə	ˈpaɾ.ˈβaʃə	ˈpaɾ.ˈβajʃa	ˈpaɾ.ˈβaʃə
	gust nou 'new taste'	ˈɡuz.ˈnɔw	ˈɡuz.ˈnɔw	ˈɡuz.ˈnɔw	ˈɡuz.ˈnɔw
(d)	*cranc petit* 'small crab'	ˈkɾaŋ.pə.ˈtit	ˈkɾaŋ.pə.ˈtit	ˈkɾaŋ.pe.ˈtit	ˈkɾajm.pə.ˈtit
	carn freda 'cold meat'	ˈkarn.ˈfɾɛðə	ˈkarɱ.ˈfɾɛðə	ˈkarɱ.ˈfreða	ˈcaɱ.ˈfɾəðə

The constraint that becomes evident for the first time here is the pre-onset coda cluster constraint *CC]$_\sigma$C (24). The motivation is fundamentally perceptual. There are inadequate cues to the place and/or manner of a consonant, especially

a stop, between two consonants. The context is somewhat less favourable than in a phrase-final CCC cluster. Utterance-final clusters of $CCC]_{\sigma}$ form are likely to be pronounced more emphatically, since utterance-final position is also that of nuclear stress, so there is likely to be more time for the speaker to achieve articulatory targets, and final position also allows more time for perception of the relatively weakly cued elements than in phrase-internal position. Thus, I argue, the ranking $*CC]_{\sigma}C$ » $*CCC]_{\sigma}$ is motivated in the production and perception of phonological phrases, as Jiménez (1999: 207) also observes.

(24) $*CC]_{\sigma}C$: A pre-onset coda does not have two consonants. (Or: a CCC cluster is not divided $CC]_{\sigma}C$.)

The more formal pronunciation in Catalonia follows largely from the constraint hierarchy mentioned previously, which is here given in a slightly elaborated form (25), including $*CC]_{\sigma}C$ alongside the other complex coda constraints, together with some Manner faithfulness constraints mentioned previously in (9). Here too IDENTW-P is included, alongside PARADIGM UNIFORMITY. For IDENTW-P the preconsonantal output is compared with the utterance-final (citation) output given in (1).

(25) (Catalonia more formal) $*C_iC_i]_{\sigma}$, IDENTSib, IDENTNasal, IDENTRhotic, IDENTLateral, IDENTPA, MAX » $*CC]_{\sigma}C$, $*3MAN]_{\sigma}$, $*CCC]_{\sigma}$ » PARADIGM UNIFORMITY, IDENTW-P, IDENTManner » $*CC]_{\sigma}$

By the constraint ranking in (25) word-final clusters are retained in words of type (23a), by IDENTPA » $*CC]_{\sigma}C$. Those in (23b) and (23c), along with the /Nk/ type (*cranc*) in (23d), are reduced by $*CC]_{\sigma}C$ » IDENTWP. The *carn* type in (23d), and similarly other homorganic clusters as in *fals* 'false', are retained by manner faithfulness IDENTSib, IDENTNasal, IDENTRhotic, IDENTLateral » $*CC]_{\sigma}C$.

What is seen in the less formal style in Catalonia is the demotion of the faithfulness constraint IDENTPA below the cluster constraint $*CC]_{\sigma}C$. This constraint, which was not relevant to the suffix data of (12), is what gives rise to the different, more reduced outcome. Thus the relevant constraint ranking for the less formal pronunciation in Catalonia is as in (26). A similar ranking of the higher constraints (27) is appropriate for Valencian, with identical cluster reduction.

(26) (Catalonia less formal) IDENTSib, IDENTNasal, IDENTRhotic, IDENTLateral, MAX » $*CC]_{\sigma}C$ » IDENTPA » $*3MAN]_{\sigma}$,$*CCC]_{\sigma}$, $*CC]_{\sigma}$, SONSEQ » PARADIGM UNIFORMITY, IDENTW-P » IDENTManner

(27) (Valencia) IDENTSib, IDENTNasal, IDENTRhotic, IDENTLateral, MAX » $*CC]_{\sigma}C$ » IDENTPA, IDENT[cont] » $*3MAN]_{\sigma}$,$*CCC]_{\sigma}$, SONSEQ » IDENTManner » $*CC]_{\sigma}$ » IDENTW-P, PARADIGM UNIFORMITY, UNIFORMITY

What is common to the hierarchies (26) and (27) is illustrated in tableau (28) with the examples *serp* 'snake', *camp* 'field', and *carn* 'meat'. MAX and the candidates that violate it are omitted, along with other constraints not relevant to these examples.

(28)

	Catalonia less formal/ Valencia *ser*$_1$*p*$_2$#C [serp]#C	IDNas	IDRho	*CC]$_\sigma$C	IDPA	IDW-P	IDMan
(a)	ser$_1$p$_2$#C			*!			
	☞ ser$_{12}$#C				*	*	*
(b)	*cam*$_1$*p*$_2$#C[kam]#C						
	kam$_1$p$_2$#C			*!		*	
	☞ kam$_{12}$#C						*
	kan$_{12}$#C				*!	*	*
	kap$_{12}$#C	*!				*	*
(c)	*car*$_1$*n*$_2$#C [karn]#C						
	☞ kar$_1$n$_2$#C			*			
	kar$_{12}$	*!				*	*
	kan$_{12}$		*!			*	*

The Majorcan version is similar in general outline, except that the only faithfulness constraints not ranked below the cluster constraint *CC]$_\sigma$C are IDENTNas and IDENTPal; the relevant hierarchy is as in (29). Note as before the high rank in Majorcan of the heterorganic marked cluster constraint *{ʃKPT}$_{[-ct]}${ʃKPT}, which means that reduced clusters ending in a non-continuant have to agree in place with the following consonant, whatever it is.

(29) (Majorca) *CC]$_\sigma$C, *3MAN]$_\sigma$, IDENTNas, IDENTPal, *{ʃKPT}$_{[-ct]}${ʃKPT}, MAX, IDENTSib » *CCC]$_\sigma$, IDENTMan » SONSEQ » *CC]$_\sigma$ » IDENTPA, IDENTW-P, PU, UNIFORMITY, INTEGRITY

For comparison with the other varieties, (30) shows the evaluation of preconsonantal *camp* and *carn* in Majorca. In this context these words are homophones.

(30)

	Majorca *cam₁p₂*#C [camp]#C	*CC]$_\sigma$C	IDNas	*{ʃKPT}$_{[-ct]}${ʃKPT}	IDMan	IDPA	IDW-P
(a)	i. cam$_1$p$_2$#C	*!					
	ii. cam$_{12}$#C			*!	*		*
	iii. cap$_{12}$#C		*!		*		*
	iv. ☞ caN$_{12}$#C				*	*	*
	v. caT$_{12}$#C		*!		*	*	*
(b)	*car₁n₂*#C [caɾn]#C						
	i. caɾ$_1$n$_2$#C	*!					
	ii. caɾ$_{12}$#C		*!		*		*
	iii. ☞ caN$_{12}$#C				*		*

Candidate (30a.ii) is, of course, only penalized before a *heterorganic* following consonant; if the following word happened to begin with a labial, *{ʃKPT}$_{[-ct]}${ʃKPT} would not be violated. The winners in (30) are represented with 'archiphoneme' [N] as a reminder that a successful output must agree in place with a following consonant. Of course, the preconsonantal clusters in the other varieties, too, are subject to heterorganic cluster constraints, but as these are rather less radical than in Majorcan, I have not drawn attention to this aspect in general.

7.6 CLUSTER REDUCTION IN INFLECTED STEMS FOLLOWED BY VOWEL-INITIAL WORDS

Stems with plural inflection have been illustrated in (12) and discussed in §7.3. When such words stand before a vowel-initial word, resyllabification takes place and a word-final consonant becomes an onset. The realizations are set out in (31).

(31)

Plural words in (12) followed by vowel	Catalonia more formal	Catalonia less formal	Valencia	Majorca
(a) *arcs amples* 'wide arches'	ˈarg.ˈzam.pləs	ˈar.ˈzam.pləs	ˈaɾ.ˈzam.ples	ˈaɾ.ˈzam.pləs
serps àgils 'quick snakes'	ˈserb.ˈza.ʒils	ˈser.ˈza.ʒils	ˈser.ˈza.d͡ʒils	ˈser.ˈza.ʒils
golfs amples 'wide gulfs'	ˈgolv.ˈzam.pləs	ˈgol.ˈzam.pləs	ˈgol.ˈzam.ples	ˈgol.ˈzam.pləs
(b) *ponts amples* 'wide bridges'	ˈpɔn.ˈzam.pləs	ˈpɔn.ˈzam.pləs	ˈpɔn.ˈzam.ples	ˈpɔn.ˈzam.pləs
camps amplis 'broad fields'	ˈkam.ˈzam.plis	ˈkam.ˈzam.plis	ˈkam.ˈzam.plis	ˈcan.ˈzam.plis

Plural words in (12) followed by vowel	Catalonia more formal	Catalonia less formal	Valencia	Majorca
alts i grossos 'tall.M.PL and large'	ˈal.zi.ˈɣɾɔsus	ˈal.zi.ˈɣɾɔsus	ˈal.zi.ˈɣɾɔsos	ˈal.zi.ˈɣɾɔsos
(c) *parts iguals* 'equal parts'	ˈpar.zi.ˈɣwals	ˈpar.zi.ˈɣwals	ˈpar.zi.ˈɣwals	ˈpar.zi.ˈɣwals
gusts/gustos agres 'sour tastes'	ˈgus.tu.ˈza.ɣɾəs	ˈgus.tu.ˈza.ɣɾəs	ˈgu.ˈza.ɣres ~ ˈgus.to. ˈza.ɣres	ˈgud.d͡za.ɣɾəs
(d) *crancs artificials* 'artificial crabs'	ˈkraŋ.zər.ti.fi. ˈsjals	ˈkraŋ.zər. ti.fi.ˈsjals	ˈkraŋ.zar.ti.fi. ˈsjals	ˈkrajn.zər. ti.fi.si.ˈals
carns assecades 'dried meats'	ˈkarn.zə.sə. ˈka.ðəs	ˈkarn.zə.sə. ˈka.ðəs	ˈkarn.za.se. ˈka.ðes	ˈcan.zə.sə. ˈca.ðəs

The realization of word-final prevocalic clusters in the more formal style of Catalonia reflects no principles beyond those that have already been considered. Resyllabification, as explained in §3.1.1, follows from ONSET outranking SYLLABIC UNIFORMITY (UNIFORMσ). The voicing of final clusters evident in (31a) follows from LAZYSIB (§5.5 (26)), which requires sibilants to be voiced before word-initial vowels, and AGREE(voice) (§5.4 (23)), which requires a coda obstruent to agree in voice with a following consonant.

The other varieties evidence, as shown also in §7.2 and §7.4, the high position of IDENTNasal, IDENTRhotic, and *CC]$_\sigma$C, which drive cluster reduction in most cases except the *carns* type, where reduction would violate IDENTNasal or IDENTRhotic. The output forms in (31a) violate IDENTWord-Phrase, as well as IDENTPA and IDENT[cont]. One might ask: why is [ˈaɾ.ˈɣzamples] not an acceptable realization of *arcs amples* in Valencian, given the dominant position of *CC]$_\sigma$C? Two constraints in principle disfavour such a realization: UNIFORMσ, and MINIMUM SONORITY DISTANCE (MSD; §3.1 (6)). From the perspective of UNIFORMσ, [ˈaɾ.ˈɣz|amples] is worse than [ˈaɾ.ˈz|amples] because in the former two segments are in the 'wrong' syllable, while in the latter only one is. From the perspective of MINIMUM SONORITY DISTANCE, [ˈaɾ.ˈɣzamples] has an onset sequence [ɣz] in which the second element is not sufficiently more sonorous than the first. It seems UNIFORMσ is the constraint required here, since it and not MINIMUM SONORITY DISTANCE penalizes, for the example *parts iguals* in (31c), the realization *[ˈpar.d͡z|i.ˈɣwals]. A unit affricate does not incur a MINIMUM SONORITY DISTANCE violation, and, of course, affricates are well-formed in Catalan onsets in general. The Syllable Contact law (SYLCON), however, does make an appearance in the realization of these forms, penalizing stop codas before fricative onsets. The relevant hierarchy for Valencia is as in (32).

(32) (Valencia) SylCon » *CC]$_\sigma$C, Uniformσ » Max, Contiguity, IdentPA, Ident[cont] » SonSeq, IdentW-P[8]

For the demonstrative tableau I choose some Valencian forms from (31a) and (31c), where the pre-vocalic realization deviates from the citation form of the plural words. The cases are *arcs*, *parts* (33) and *gusts* (34).

(33)

	Valencia *arcs*#V /aɾ$_1$k$_2$+z/ [aɾks]#V	SylCon	IdRhot	*CC]$_\sigma$C	Uniσ	Max	IdPA	Id[cont]	IdW-P
(a)	aɾg.z	*!		*	*				
	aɾ$_1$.ɣ$_2$z				**!			*	*
	☞ aɾ$_1$.z				*	*			*
	☞ aɾ$_{12}$.z				*		*		*
	ag$_{12}$.z	*!	*		*		*		*
(b)	*parts*#V /paɾ$_1$t$_2$+z$_3$/ [paɾ$_1$t͡s$_{23}$]#V								
	paɾ$_1$d$_2$.z$_3$	*!		*	*				
	paɾ$_1$.d͡z$_{23}$				**!				
	paɾ$_1$.ð$_2$z$_3$				**!			*	*
	☞ paɾ$_1$.z$_3$				*	*			*
	☞ paɾ$_{12}$.z$_3$				*			*	*

(Valencia) SylCon, IdentRhot » *CC]$_\sigma$C, Uniformσ » Max, IdentPA, Ident[cont], IdentW-P

In (33) the relatively faithful forms without reduction (top two candidates in (33a), the top three in (33b)) fall prey to *CC]$_\sigma$C or Uniformσ. The remaining candidates are worse from the point of view of IdentW-P but have good syllable contacts with minimum Uniformσ violation (only that necessary to fulfil Onset). The tied winners in each case are phonetically indistinguishable.

The interpretation of evaluation by Uniformσ in (34) involves checking correspondences rather than output segments, so that *[gu.d͡zV] is a worse match for [gusts.] than [gu.zV] is, despite each having only one segment in the final onset, on the grounds that *[gu.d͡zV] has more coda material transferred to an onset. A similar observation led to preferring *3Man]$_\sigma$ to *CCC]$_\sigma$ as an active cluster constraint. Otherwise, it is difficult to see on what other grounds reduced [gu.zV] should be preferred to *[guz.d͡z], which is more faithful as well as being syllabically well-formed.

[8] SylCon » Uniformσ is demonstrated by pre-vocalic *tots* 'all.M.pl': [tod.d͡zV] ≻ *[tod.zV].

(34)

	Valencia *gusts*#V /gus$_1$t$_2$+z$_3$/ [gus$_1$t͡s$_{23}$]#V	*GEMSib	IDSib	SYLCON	UNIσ	CONTIG	MAX	ID[cont]	IDW-P	UNIF
(a)	guz$_1$d$_2$.z$_3$			*!	*				*	
(b)	guz$_1$.d͡z$_{23}$				**!					*
(c)	guz$_1$.z$_3$	*!			*	*	*		*	
(d)	guz$_{12}$.z$_3$	*!			*			*	*	*
(e)	gu.z$_3$				*	*!	**		*	
(f)	☞ gu.z$_1$				*		**		*	
(g)	gu.z$_{123}$				**!*			*	*	**
(h)	gud$_{12}$.z$_3$		*!		*			*	*	*
(i)	gu.d͡z$_{23}$				**!	*	*		*	*
(j)	gud$_1$.d͡z$_{23}$				**!			*	*	*
(k)	gu.d͡z$_{123}$				**!*				*	**

(Valencia) *GEMSib, IDENTSib, SYLCON » UNIFORMσ » CONTIG, MAX, IDENT [cont], IDENTW-P, UNIFORMITY

The Majorcan realization is similar in most respects, except in the case of *gusts agres* (35). Where Valencian has reduction of /guzdz/ to [guz] because UNIFORMσ dominates IDENTMan or IDENTW-PMan (and other relevant faithfulness constraints), Majorcan has [gud.d͡z], that is, with a geminate affricate, which is the

(35)

	Majorca *gusts*#V /gu$_1$s$_2$t$_3$+z$_4$/ [gu^1$_1$t͡s^{2,3,4}$_2$]#V	SYLCON	*GEMSIB	*CC]$_\sigma$C	IDW-PMan	IDMan	UNIσ	CONTIG	INTEG
(a)	gu$_1$zd$_2$.z$_2$	*!		*	*		*		
(b)	gu$_1$z$_2$.z$_2$		*!		*	*	*		
(c)	gu$_1$.z$_2$				*!	*	*	*	
(d)	gu$_1$.t$_2$				*!	*		*	
(e)	gu$_1$d$_2$.z$_2$	*!				*	*	*	
(f)	☞ gu$_1$d$_2$.d͡z$_2$					*	*	*	*
(g)	gu$_1$.d͡z$_2$					*	**!	*	

(Majorca) **SYLCON**, *GEMSIB, *CC]$_\sigma$C, **IDENTW-PMan** » UNIFORMσ » CONTIGUITY, INTEGRITY

best correspondent to the citation plural form [gut͡s], given the general pattern in the language for gemination of intervocalic affricates (see §2.1.3.3). That is, in Majorca IDENTMan or IDENTW-PMan dominates UNIFORMσ. In (35), which compares candidates for the prevocalic plural *gusts* 'tastes' in Majorcan, candidate (35a) *[$gu_1zd_2.z_2$], though relatively faithful to the input, in addition violates CONTIGUITYW-P, in that it contains a case of [z] without a correspondent in the citation form [gut͡s] of the plural word. Other candidates with CONTIGUITYW-P violations are not shown.

The representation [$gu^1{}_1\widehat{ts}^{2,3,4}{}_2$]#V in the heading of (35) indicates I-O correspondents with superscripts, and the basis for Word-Phrase correspondence with subscripts. In the tableau the I-O correspondence is not separately indicated, though it is evaluated by CONTIGUITY. The winning candidate (35f) is understood to beat (35g) on UNIFORMσ in that, whereas (35g) transfers all of the coda material (and correspondence) of the plural citation form to the pre-vocalic onset, (35f) retains something in the coda despite the prevocalic onsets of (35f) and (35g) being identical. There might be a case for invoking MAXμ here, though the role of such a constraint in Catalan is otherwise not clear. In evaluating outputs against the word form [gut͡s], the Word-Phrase correspondence version of IDENTManner is invoked, standing above UNIFORMσ; forms lacking a stop or a sibilant lose.

7.7 CLUSTER REDUCTION IN INFLECTED STEMS FOLLOWED BY CONSONANT-INITIAL WORDS

The final pattern involving cluster reduction is the one potentially involving four-consonant sequences: a stem cluster followed by inflectional /+z/, followed by a consonant initial word. They are illustrated in (36).

(36)	Plural words in (12) followed by consonant	Catalonia more formal	Catalonia less formal	Valencia	Majorca
	(a) *arcs baixos*	ˈargz.	ˈarz.	ˈarz.	ˈar.
	'low arches'	ˈβaʃus	ˈβaʃus	ˈβajʃos	ˈβaʃos
	serps grans	ˈserbz.	ˈserz.	ˈserz.	ˈser.
	'large snakes'	ˈɣrans	ˈɣrans	ˈɣrans	ˈɣrans
	golfs grans	ˈgolvz.	ˈgolz.	ˈgolz.	ˈgol.
	'large gulfs'	ˈɣranz	ˈɣrans	ˈɣrans	ˈɣrans
	(b) *ponts baixos*	ˈpɔnz.	ˈpɔnz.	pɔnz.	ˈpɔm.
	'low bridges'	ˈβaʃus	ˈβaʃus	ˈβajʃos	ˈbaʃos
	camps nous	ˈkamz.	ˈkamz.	kamz.	can.
	'new fields'	ˈnɔws	ˈnɔws	ˈnɔws	ˈnɔws
	alts càrrecs	ˈals.	ˈals.	ˈals.	ˈal.ˈca.rets
	'high positions'	ˈka.rəks	ˈka.rəks	ˈka.reks	

Plural words in (12) followed by consonant	Catalonia more formal	Catalonia less formal	Valencia	Majorca
(c) *parts baixes* 'lower parts'	ˈparz. ˈβaʃəs	ˈparz. ˈβaʃəs	ˈparz. ˈβajʃes	ˈpar. ˈβaʃəs
gusts/gustos nous 'new tastes'	ˈgus.tuz. ˈnɔws	ˈgus.tuz. ˈnɔws	ˈguz. ˈnɔws ∼ gus.toz. ˈnɔws	ˈgun.ˈnɔws
(d) *crancs petits* 'small crabs'	ˈkraŋs.pə. ˈtits	ˈkraŋs.pə. ˈtits	ˈkraŋs. pe.ˈtits	ˈkrajm. pə.ˈtits
carns fredes 'cold meats'	karns. ˈfrɛðəs	karns. ˈfrɛðəs	karns. ˈfreðes	ˈcaɱ. ˈfrəðəs

The more formal pronunciation of Catalonia shows the citation forms of the plural words with minimal adaptation, namely voicing assimilation. Less formal style in Catalonia, and Valencian pronunciation, show essentially the same forms before consonants as before vowels. These phrase-internal forms show reduction of cluster obstruents which are unreduced in the citation plural forms. This phrase-internal coda reduction shows the activity of either a *CCC]$_\sigma$C internal coda markedness constraint, or the related *3MANNER]$_\sigma$C constraint. These forms also show, of course, that several faithfulness constraints, for example MAXSib, IDENTSib, MAXRhotic, IDENTRhotic, MAXNasal, and IDENTNasal, stand above *CC]$_\sigma$C or *2MAN]$_\sigma$C. Majorcan shows further reduction over the citation and pre-vocalic plurals, in which stem-final obstruents are absent (except *gusts* with obstruent affricate [guts͡] ∼ [gud.d͡z]). This further reduction is in line with the Majorcan strong 'contact' coda constraint *CC]$_\sigma$C. In fact, it was shown above in connection with *CCC]$_\sigma$ that rather than counting segments in complex cluster constraints, it seems to be more relevant to count manners of articulation, so the active constraint was *3MANNER]$_\sigma$. The same point is relevant here, and the active constraint is *2MANNER]$_\sigma$C (see below). The motivation for these two coda manner constraints is the same (as indeed it is with Complex Coda constraints in general), namely the articulatory difficulty of changing manners of articulation more than once within a cluster, where time is limited and the effectiveness of perceptual cues is compromised. But not all three-consonant clusters are bad: a single coda consonant followed by a complex onset is not penalized, as in *ombra* [ˈom.bɾə] 'shadow'. Here the sonorants have good inherent cues, and the medial stop has adequate perceptual cues. That is, it is not surprising that we do not see active a simple *CCC constraint that does not mention syllable organization.

More formal pronunciation of Catalonia shows faithfulness constraints, IDENTNasal, IDENTSib, IDENTPA, MAX, dominating cluster constraints *CCC]$_\sigma$, or *3MAN]$_\sigma$C, or *3MAN]$_\sigma$. Less formal pronunciation in Catalonia, and Valencian pronunciation, show largely the same constraints active as were seen in

pre-vocalic environments, that is to say, IDENTSib, IDENTNasal, IDENTRhotic » *CC]$_\sigma$C » MAX, IDENTPA, IDENT[cont], etc.

Tableau (37) illustrates the reduction of pre-consonantal *gusts* 'tastes' to [gus] in Valencian. Whereas in most of the examples in (36) coalescence of sonorant + sibilant clusters is not a permitted repair to potential *2MANNER]$_\sigma$C violations, showing that IDENTNasal, IDENTRhotic, IDENTLateral, IDENTSibilant dominate *2MANNER]$_\sigma$C, the situation with an /sts/ cluster is different.

(37)

	Valencia *gusts*#C /gus$_1$t$_2$+z$_3$/ [gus$_1$t͡s$_{23}$]#C	*GEMSib	IDSib	MAXSib	*CC]$_\sigma$C	*2MAN]$_\sigma$C	CONTIG	MAX	ID[cont]	IDW-P	UNIF
(a)	gus$_1$t͡s$_{23}$#C				*!	*					*
(b)	gus$_1$s$_3$#C	*!			*			*		*	
(c)	gus$_{12}$s$_3$#C	*!			*				*	*	*
(d)	gus$_3$#C			*!			*	**		*	
(e)	gus$_1$#C			*!				**		*	
(f)	gut$_{123}$#C		*!						*	*	**
(g)	☞ gus$_{123}$#C								*	*	**
(h)	gut͡s$_{23}$#C			*!		*	*	*		*	*
(i)	gut͡s$_{123}$#C					*!			*	*	**

Tableau (37) shows how, while sibilant deletion is ruled out (37d, 37e, 37h), coalescence remains an option where the outcome is a sibilant (37g, 37i). Of these, however, the form with an affricate (37i) falls to *2MANNER]$_\sigma$C, leaving the candidate with multiple coalescence [gus$_{123}$] as the winner.

In Majorca the high ranking of the preconsonantal coda constraint *CC]$_\sigma$C means that in phrases, as in words, internal codas have no more than one consonant. A notable effect of this is that the plural morph /+z/ is obliterated in consonant-final words standing before another consonant, as Dols (2000: 337) points out. So *arcs* in *arcs baixos* 'low arches' [ˈaɾ.ˈβaʃos] sounds just like *arc* in *arc baix* 'low arch' [ˈaɾ.ˈβaʃ]. There is one striking difference between singular and plural that is apparent in comparing (36) with (24). Pre-consonantal *gust* 'taste' is realized as [gus] whereas pre-consonantal *gusts* 'tastes' is realized as [guT], that is to say, the plural form ends in a 'stop' that agrees in place and manner with a following consonant. ('Stop' in scare quotes because like underlying obstruent stops it is realized as an underspecified consonant in pre-consonantal environments.) This difference between the realization of *gust* and *gusts* provides substantial evidence for the constraint hierarchy of Majorcan, and also in support of the suggestion made just above that the active pre-consonantal coda cluster constraint is actually *2MAN]$_\sigma$C. It is this constraint

that, for *gusts*#C, prefers the candidate [guT#C] over [gut͡s#C]. Each of these has the form CVC.C, but the affricate in [gut͡s] falls foul of *2MAN]$_\sigma$C. Otherwise [guT] would lose on IDENTSib, or on REALIZEMorph, and might only win on UNIFORMITY, in that [gut͡s]'s affricate displays at least one case of fusion. (And for this to work [guT] would need to have fewer fusions, which means it would lose by MAX, having two deletions.) The difference between the singular and plural realization of *gusts* in pre-consonantal position in Majorcan is in fact 'inherited' from the difference between the citation forms of the singular *gust* [gust] 'taste' and the plural *gusts* [gut͡s] 'tastes'. Tableaux (38) and (39) compare the realization of pre-consonantal *gust* [gus] and *gusts* [guT] in Majorcan.

(38)

	Majorca *gust*#C /gus$_1$t$_2$/ [gus$_1$t$_2$]#C	*2MAN]$_\sigma$C	*{ʃKPT}$_{[-ct]}${ʃKPT}	IDW-P[–cont]	IDSib	MAX	CONTIGW-P	IDW-PMan	IDMan	CONTIG	IDPA
(a)	gus$_1$t$_2$#C	*!	*								
(b)	gus$_{12}$#C			*!				*	*		
(c)	☞ gus$_1$#C					*					
(d)	gut$_{12}$#C		*!		*			*	*		
(e)	guT$_2$#C					*	*!			*	
(f)	guT$_{12}$#C				*!			*	*		*

The faithful candidate (38a) violates the coda cluster constraint *2MAN]$_\sigma$C, highly ranked in Majorcan. In contrast to what appeared in (37), in this contest the fusion candidates (38b, 38d, 38f) lose, falling to IDENTW-P[–cont] (which requires corresponding elements in citation forms and their contextual realizations to agree in the feature [–cont]) or to IDENTSib. (Candidate (38d) is a version of (38f), included as a reminder that in Majorcan any candidate with a heterorganic obstruent coda will lose; the winner must display place assimilation.) Of the deletion candidates, (38e) *guT$_2$#C has a contiguity violation—both Word–Phrase CONTIGUITY and I-O CONTIGUITY are included in the tableau, though their activity is the same in this case—so (38c) wins.

(39)

	Majorca *gusts*#C /gus^{1}t^{2}+z^{3}/ [gut͡s^{123}]#C	*GemSib	*2Man]$_{\sigma}$C	Realm	IdW-P[–cont]	IdSib	Max	ContigW-P	IdMan	Contig	IdPA
(a)	gus^{1}t͡s^{23}#C		*!					*			
(b)	gus^{1}s^{23}#C	*!			*			*	*		
(c)	gus^{123}#C				*!				*		
(d)	gus^{1}#C			*!	*		**				
(e)	guT^{12}#C			*!		*	*		*		
(f)	guT^{2}#C			*!			**			*	
(g)	guT^{23}#C					*	*!		*	*	
(h)	☞ guT^{123}#C					*			*		
(i)	gut͡s^{23}#C		*!				*			*	
(j)	gut͡s^{123}#C		*!								

Tableau (39) shows candidates for pre-consonantal plural *gusts* 'tastes'. Raised numerals indicate I-O correspondence. The faithful candidate, (39a), has an over-complex coda, as well as displaying a mismatch with the isolation word form (here penalized via ContiguityWP, as *[gust͡s] inserts a consonant absent in [gut͡s]). Candidates (39i, 39j), pronounced like the isolation form, also have an over-complex pre-consonantal coda. Candidate (39b) is better from the perspective of *2Manner]$_{\sigma}$C, but displays an ungrammatical geminate sibilant. Candidate (39d) is the same as the winning candidate in (38), but in the present case falls to RealizeMorph as the form lacks a correspondent of the plural /+z/. A form (39c) pronounced the same as (39d), but with coalescence rather than deletion, passes RealizeMorph, but is eliminated by IdentW-P[–cont]. Of the [guT] forms (39e–39h) the winner is the one with the most coalescence. It is interesting to observe, in the comparison of (38) with (39), that a morphophonemic alternation that seems to display some opacity (/gust/→ [gus]/__#C, /gust+z/ → [guT]/__#C) actually follows from an orthodox constraint hierarchy.

7.8 SUMMARY OF CONSTRAINT RANKINGS RELATING TO WORD-FINAL CLUSTERS

Catalonia more formal, cumulative hierarchy: ONSET, *GEMSIB, $*C_iC_i]_\sigma$, OPDEPV[9], IDENTNAS, IDENTRHOTIC, IDENTLATERAL, IDENTSIB, IDENTPA, MAX » $*CC]_\sigma C$, *3MAN$]_\sigma$, $*CCC]_\sigma$, SONSEQ, SYLCON » IDENTW-P » IDENTMAN » $*CC]_\sigma$ » UNIFORMσ, PU, CONTIGUITY, UNIFORMITY, INTEGRITY.

Catalonia less formal, cumulative hierarchy: ONSET, *GEMSIB, $*C_iC_i]_\sigma$, OPDEPV, IDENTNAS, IDENTRHOTIC, IDENTLATERAL, IDENTSIB, MAX » $*CC]_\sigma C$ » IDENTPA » *3MAN$]_\sigma$, $*CCC]_\sigma$, $*CC]_\sigma$, SONSEQ, SYLCON » IDENTW-P » IDENTMAN, UNIFORMσ, PU, CONTIGUITY, UNIFORMITY, INTEGRITY.

Valencia cumulative hierarchy: ONSET, SYLCON, *GEMSIB, OPDEPV, IDENTNAS, IDENTRHOTIC, IDENTLATERAL, IDENTSIB » UNIFORMσ » MAXSIB » *3MAN$]_\sigma$C » *2MAN$]_\sigma$C » MAX, IDENTPA, IDENT[cont] » *3MAN$]_\sigma$, $*CCC]_\sigma$, SONSEQ, IDENTMAN » $*CC]_\sigma$ » IDENTW-P, PU, CONTIGUITY, UNIFORMITY, INTEGRITY.

Majorca cumulative hierarchy: ONSET, SYLCON, *GEMSIB, $*CC]_\sigma C$, *3MAN$]_\sigma$, *2MAN$]_\sigma$C, $*\{\int KPT\}_{[-ct]}\{\int KPT\}$, OPDEPV, IDENTNAS, IDENTLATERAL » IDENTRHOTIC, IDENTPAL, REALIZEMORPH » IDENTW-P[-cont], IDENTSIB » MAX, CONTIGW-P » IDENTW-PMAN » $*CCC]_\sigma$, IDENTMAN » SONSEQ, $*CC]_\sigma$, UNIFORMσ » IDENTPA, IDENTW-P, PU, CONTIGUITY, UNIFORMITY, INTEGRITY.

Common: ONSET, *GEMSIB, OPDepV, IDENTNAS » SONSEQ (») $*CC]_\sigma$ » PU, CONTIGUITY, UNIFORMITY, INTEGRITY.

[9] On OPDEPV, see §8.6.

8

EPENTHESIS AND OTHER SONORITY-RELATED PHENOMENA

8.1 INTRODUCTION

I take up in this section some phenomena that illustrate further the three sonority constraints introduced in §3.1: the Sonority Sequence principle (SonSeq), the Minimum Sonority Distance principle (MSD), and the Syllable Contact law (SylCon). The major focus in this chapter is vowel epenthesis, that well-known process by which 'underlying' sequences that violate the Sonority Sequence principle (SonSeq) are made acceptable by the 'insertion' of a default or unspecified vowel. Thus, for example, the infinitive of the verb 'write' *escriure* consists of the stem /ˈskRiw/[1] followed by the consonantal infinitive inflection /+ɾ/. The output is not *[ˈskɾiwɾ] but [əsˈkɾiwɾə], in eastern Catalan pronunciation, with epenthesis of [ə] in both initial and final position, to allow proper syllabification of initial [sC] and final [wɾ]. (Western Catalan has [esˈkɾiwɾe] or [asˈkɾiwɾe].) There are accounts of the Catalan epenthesis phenomena from the perspective of Generative Phonology in Mascaró (1976/8), Wheeler (1975/9: 11–32; 1987), and DeCesaris (1987). Lloret (2002a) provides a comprehensive description. While it is true that there are no output forms with *[sC-] initials, nor -C_1C_2 finals except where C_1 is more sonorous than C_2 (and not all such finals are admissible, as *escriure* shows), it does not necessarily follow that a process of 'epenthesis' takes place. All that one may directly infer is that contrast between zero and the presence of the default vowel is neutralized in such contexts. Given 'Richness of the Base', the relevant constraints would ensure that, for example, both /sC-/ and /VsC-/ inputs were realized [əsC-].

There is a small amount of morphological evidence for true /sC-/ roots. Some representative examples are given in (1).[2] In contrast with the data in (1), many roots beginning with [əsC-] retain [ə] after a consonant-final prefix or pre-bound root, as in *estable* 'stable' – *inestable* 'unstable'; *estructura* 'structure' – *superestructura* 'superstructure'; *esperar* 'hope.INF' – *desesperar* 'despair.INF'. Such contrast implies that the roots in (1), and some others like them, do actually have

[1] I use a 'surfacey' representation of underlying high vocoids /[u] ∼ [w]/ and /[i] ∼ [j]/ in this chapter.

[2] And the stress pattern of the verb *estar* makes sense if the root 'really' begins /st-/, as explained in §9.4.

/sC-/ initials, and illustrate the ranking SonSeq » DepV. However, the apparent contrast between, for example, *superstrat* /supeR+ˈstrat/ 'superstratum' (initial #sC- in the root) and *superestructura* /super+estrukˈtuɾ+a/ (initial #esC- or #VsC- in the root) may well be an artefact of inconsistent language standardization. Such, anyway, is the view of Cabré (1993a), for whom the only genuine options are, for *superstrat*, (a) interpreted as morphologically complex: /supeR-+eˈstrat/, pronounced [supəɾəsˈtɾat], or (b) interpreted as underived /supeR-ˈstrat/, pronounced [supərsˈtrat]. Nearly all the 'derived' forms in (1) are synchronically simple, according to this view, and do not contain examples of root-initial /sC-/.

(1) Possible /sC-/ roots	Derived forms
espirar [əspiˈɾa] /ˈspiR+ˈa+ɾ/ 'breathe.INF, blow.INF'	*inspirar* [inspiˈɾa] /in+/ 'breathe-in.INF'; *transpirar* [trənspiˈɾa] /tRans+/ 'transpire.INF'
estrat [əsˈtɾat] /ˈstRat/ 'stratum'	*substrat* [supsˈtɾat] /sup+/ 'substratum'; *superstrat* [supersˈtrat] /supeR+/ 'superstratum'
estrènyer [əsˈtɾɛɲə] /ˈstRɛɲ+ɾ/ 'squeeze.INF'	*constrènyer* [kunsˈtɾɛɲə] /kon+/ 'constrain.INF'
espuma [əsˈpumə] /ˈspum+a/ 'scum'	*despumar* [dəspuˈma] /des+spum+a+ɾ/ 'remove.INF scum from'
escriure /ˈskRiu+ɾ/ 'write.INF'	*circumscriure* [ˌsirkumsˈkɾiwɾə] /siRkum+/ 'circumscribe.INF'; *inscriure* [insˈkɾiwɾə] /in+/ 'inscribe.INF'; *transcriure* [trənsˈkɾiwɾə] /tRans+/ 'transcribe.INF'

What of final 'epenthesis'? Can one say that the pair *teatre* [teˈatrə] – *teatral* [teəˈtɾal] actually shows epenthesis in the base (/teˈatɾ/) rather than truncation of /teˈatɾe/ before the vowel-initial suffix /+ˈal/? Actually the evidence is indeterminate. What can be said, at least, is that *[teˈatɾ] is not a possible output, nor is *[teatɾeˈal] a possible output of a base like /teatRe/ followed by the suffix /+ˈal/. Final /e/ of a base ([ə] or [e] according to dialect) is never retained before a vowel-initial suffix, even when it is not 'epenthetic' after an ill-formed coda (as in *mode* [ˈmɔðə] 'mood' – *modal* [muˈðal] 'modal' *[muðeˈal]). There are a few examples of surface [-eˈV] etc. involving a vowel initial suffix, such as *areal* [əɾeˈal] 'areal' corresponding to the base *àrea* [ˈareə], or *traqueïtis* [tɾəkeˈitis] 'tracheitis' corresponding to *tràquea* [ˈtɾakeə] 'trachea'. And such bases themselves might be subject to truncation if their structure were such as /ˈare+a/, for *àrea*, with /+a/ realizing the feminine gender inflection. Before feminine /+a/ 'truncation' is indeed expected, as in *aspre* 'rough.M' – *aspra* 'rough.F', which are [ˈaspɾe] and [ˈaspra] not *[ˈasprea], respectively, in western Catalan. (In

eastern Catalan *aspre* 'rough.M' and *aspra* 'rough.F' are homophonous: [ˈasprə].) Since in words like *àrea* there is no independent evidence for a root /ˈare/, 'truncation' would be expected to be blocked by faithfulness constraints to the base /ˈarea/, such as DEPC, MAXV, UNIFORMITY, or (in particular) CONTIGUITY, standing above whatever anti-hiatus markedness constraint, such as ONSET, drives truncation. Of course the final /a/ of the base /ˈarea/ would itself be subject to truncation before /+al/, since edge elements are not protected by CONTIGUITY. Pursuing this further requires more morphological investigation than there is scope for in this book. The analysis I offer here seeks to capture the observations mentioned above: that *[teˈatɾ] and other SONSEQ violations are not possible outputs, nor is *[teatɾeˈal] a possible output of a base like /teatɾe/ followed by the suffix /+ˈal/. I return shortly to the question why, for *escriure* 'write.INF', the realization *[əsˈkɾiwr] is ill-formed, despite its coda [wr] conforming to the Sonority Sequence principle—a glide being more sonorous than a liquid.

8.2 NON-EDGE EPENTHESIS

In addition to initial and final 'epenthesis' driven by SONSEQ, Catalan also manifests some word-internal epenthesis. All the cases occur in verb morphology, specifically in verbs of conjugation II, where consonant-final stems are followed by the consonantal inflexion /+ɾ/, in the infinitive, and the future and conditional paradigms. The data in (2) takes representative forms (1PL.PR.IND) showing the shape of the stem before a vowel-initial affix, and a second set (1PL.FUT.IND) in which the same stems are followed by /+ɾ/. (2a) illustrates stem-final consonants that do not provoke epenthesis before /+ɾ/; (2b) illustrates epenthesis of [d]; and (2c) illustrates epenthesis of the default vowel /e/ (western [e], eastern [ə]). Central Catalan pronunciations only are given.

The forms without epenthesis (2a) belong to stems ending in non-sibilant obstruents. Those in (2b) with epenthesis of [d] have stems ending in an alveolar nasal or lateral. Epenthesis 'repairs' the cluster */nɾ/ or */lɾ/. Those in (2c) with epenthesis of [ə] have stems ending in a non-alveolar nasal, or in a sibilant. Epenthesis 'repairs' the clusters */ɲɾ/ */mɾ/, */ʃɾ/, */sɾ/. The (2d) example most plausibly illustrates stem-final epenthesis after a cluster /RR/ which cannot be syllabified as a coda; note that /koRR/ is realized [ˈkorə]/ [kurə] before any affix consonant (*corres* /ˈkoRR+z/ 'run.2SG.PR.IND', *corregués* /ˈkoRR+g+ˈes/ 'run.3SG.PST.SUBJ'), and also word-finally (*corre* /ˈkoRR/ 'run.3SG.PR.IND'). None of the unrepaired clusters in (2a)-(2c) is acceptable in an output, in any circumstances. The questions are: why not? and why are these repairs preferred to other conceivable ones? Answering these questions

invokes the two more specific sonority-related constraints, SylCon and MSD (Minimum Sonority Distance), which were introduced in §3.1. Taking *nɾ as a representative case, the syllabification [n.ɾ] violates SylCon because coda [n] is less sonorous than onset [ɾ]. But the syllabification [.nɾV] violates MSD, because, while [.nɾV] respects the Sonority Sequence principle, the difference in sonority between [n] and [ɾ] is insufficient to make an adequate onset. The faithfulness constraint Dep is ranked below both of the sonority markedness constraints.

(2) Verb roots before +V suffix		Verb roots before /+ɾ/	
(a) *perdem* 'we lose'		*perdrem* 'we shall lose'	
/ˈpɛRd+ˈɛ+m/	pəɾˈðɛm	/ˈpɛRd+ɾ+ˈɛ+m/	pəɾˈðɾɛm
admetem 'we admit'		*admetrem* 'we shall admit'	
/amˈmɛt+ˈɛ+m/	əmməˈtɛm	/amˈmɛt+ɾ+ˈɛ+m/	əmməˈtɾɛm
(b) *volem* 'we want'		*voldrem* 'we shall want'	
/ˈbɔl+ˈɛ+m/	buˈlɛm	/ˈbɔl+ɾ+ˈɛ+m/	bulˈdɾɛm
ofenem 'we offend'		*ofendrem* 'we shall offend'	
/oˈfɛn+ˈɛ+m/	ufəˈnɛm	/oˈfɛn+ɾ+ˈɛ+m/	ufənˈdɾɛm
(c) *pertanyem* 'we belong'		*pertanyerem* 'we shall belong'	
/peRˈtaɲ+ˈɛ+m/	pərtəˈɲɛm	/peRˈtaɲ+ɾ+ˈɛ+m/	pərtəɲəˈɾɛm
espremem 'we squeeze'		*espremerem* 'we shall squeeze'	
/esˈpRem+ˈɛ+m/	əspɾəˈmɛm	/esˈpRem+ɾ+ˈɛ+m/	əspɾəməˈɾɛm
creixem 'we grow'		*creixerem* 'we shall grow'	
/ˈkReʃ+ˈɛ+m/	kɾəˈʃɛm	/ˈkReʃ+ɾ+ˈɛ+m/	kɾəʃəˈɾɛm
vencem 'we overcome'		*vencerem* 'we shall overcome'	
/ˈbɛns+ˈɛ+m/	bənˈsɛm	/ˈbɛns+ ɾ+ˈɛ+m/	bənsəˈɾɛm
(d) *correm* 'we run'		*correrem* 'we shall run'	
/ˈkoRR+ˈɛ+m/	kuˈrɛm	/ˈkoRR+ɾ+ˈɛ+m/	kurəˈɾɛm

Parker's account of sonority (2002) claims that the sonority hierarchy is a phonological universal that correlates with several acoustic or physical measurements, most strongly with intensity, but also (negatively) with intra-oral air pressure and with airflow, at least as far as consonants in onsets are concerned. Parker also provides experimental psycholinguistic evidence for the psychological reality of the sonority hierarchy among a population of native English speakers. He develops a sonority scale, derived from measurements of speakers of NE US English and Colombian Spanish, that he offers as universal, which I quote in (3).

(3) Sonority scale (Parker 2002: 240)

low vowels	16
mid vowels (except /ə/)	15
high vowels (except /ɨ/)	14
/ə/	13
/ɨ/	12
glides	11
/ɹ/	10
laterals	9
flaps [[ɾ]]	8
trills [[r]]	7
nasals	6
/h/	5
voiced fricatives	4
voiced stops and affricates/voiceless fricatives	3
voiceless fricatives/voiced stops and affricates	2
voiceless stops and affricates	1

Note that Parker's hierarchy contains one alternative ranking: voiced stops may count as more or less sonorant than voiceless fricatives. He claims that, apart from this, 'no language may permute (reverse) any of the 16 classes' (2002: 241), but concedes, as his English psycholinguistic experiment later shows, that /h/ may fit better elsewhere than it is placed here. Of course, it is normal for languages to group adjacent classes, and not to distinguish phonologically within the conflated classes. No doubt Parker's claim of impermutability is too strong, based as it is on data from only two languages. I mention this because there is reason from the phonological behaviour of Catalan to distinguish sibilant from non-sibilant obstruents (with sibilants more sonorous, a possibility that Parker himself discusses on the basis of the English psycholinguistic data (2002: 275–6)). There is also evidence that in Catalan /l/ is less sonorous than /ɾ/. In comparison with Colombian Spanish, from which Parker's ranking of /l/ and /ɾ/ comes, it may be that the characteristic velarized quality of Catalan /l/ contributes to its being less sonorous. The sonority scale in (3) also makes no distinction between affricates and (plosive) stops, and such a conflation is rather unsatisfactory when one's focus is the relative markedness of consonant sequences in syllable structure. It implies, for example, that [d͡zɾ-] is no more marked an onset than [dɾ-]. In deriving a sonority scale for Catalan, I 'decompose' affricates, in line with the way the syllable syntax of Catalan treats them. From this perspective, the absence of clusters like [d͡zɾ-] follows automatically from the Minimum Sonority Distance constraints.

The phonological evidence from Catalan in fact supports the sonority classification in (4). The numerical values are no more significant in themselves than Parker's are meant to be; their point is to reflect with whole integers the clustering of the sound types and 'minimum distance' as observed in complex onsets.

The sonority taxonomy in (4), with its associated relative numerical values, allows us to draw up a minimum sonority distance constraint for onsets as in (5), followed by examples of permitted and excluded onsets.

(4) Sonority ranking (Catalan)

15	9	8	5	4	2	0
Glides	Tap	Laterals, trill	Nasals	Sibilant continuants	Non-strident continuant obstruents	Non-sibilant stops/strident continuants
j w	ɾ	l ʎ r	m n ɲ	s ʃ z ʒ	β ð ɣ	f p t k b d g

(5) MSD6: In an onset sequence C_1C_2, the value of $C_2 - C_1 \geq 6$

Examples:

V.fɾV, V.pɾV, V.tɾV, V.kɾV, Vm.bɾV, Vn.dɾV, Vŋ.gɾV	$[_\sigma$0-9V
V.flV, V.plV, V.klV, Vm.blV, Vŋ.glV	$[_\sigma$0-8V
V.βɾV, V.ðɾV, V.ɣɾV	$[_\sigma$2-9V
V.βlV, V.ɣlV	$[_\sigma$2-8V
V.sjV, V.ʒjV	$[_\sigma$4-15V
V.mjV, V.nwV	$[_\sigma$5-15V
V.rjV, V.ʎwV	$[_\sigma$8-15V
V.ɾjV	$[_\sigma$9-15V
*V.slV, *V.sɾV	*$[_\sigma$4-9V
*V.nɾV	*$[_\sigma$5-9V
*V.mlV, *V.nlV	*$[_\sigma$5-8V
*V.βʒV, *V.ðzV, *V.ɣzV	*$[_\sigma$2-4V
*V.lɾV	*$[_\sigma$8-9V
V.fɾjV, V.pɾwV, V.tɾjV, V.kɾwV, Vm.bɾjV, Vn.dɾwV, Vn.gɾjV	$[_\sigma$0-9-15V
V.fljV, V.pljV, V.kljV, Vm.bljV, Vŋ.gljV	$[_\sigma$0-8-15V
V.βɾjV, V.ðɾjV, V.ɣɾjV	$[_\sigma$2-9-15V
V.βljV, V.ɣljV	$[_\sigma$2-8-15V

I am now in a position to show how internal epenthesis follows from, and is restricted by, quite well-motivated constraints. The (2b) examples with consonant epenthesis avoiding *lɾ and *nɾ need to be contrasted with the morpheme-internal clusters [l.r], as in *folro* [ˈfol.ru] 'lining', *Enric* [ən.ˈrik] 'Henry'. Recall from §2.1.6.1 that for the input rhotic /R/ the trill [r] automatically occurs onset-initially except after a vowel, tap [ɾ] elsewhere in onsets. In considering *voldrem* 'we shall want', it is necessary to show why epenthesis is permitted in words of this shape, so as to provide a better syllable contact, whereas no such epenthesis takes place in *folro*. The CONTIG(UITY) constraint, a high-ranking member of the DEP family, penalizes interruption of lexical material (Jiménez 1999: 56, after

Kenstowicz 1994). It thus favours [ˈfol.ru] over *[ˈfol.dɾu] for *folro* /folR+o/. But CONTIG is not violated by insertion of material between morphemes. In the verb forms, however, there is inter-paradigmatic pressure for the /+R/ affix (to give it a neutral representation) to be realized consistently as a tap [ɾ] (Bonet & Lloret 1998: 91–3), because it occurs in post-vocalic, or post-non-sibilant, position in the great majority of verbs (e.g. *parlarà* [pərləˈɾa] 'speak.3SG.FUT', *rebrà* [rəˈβɾa] 'receive.3SG.FUT'). This paradigm uniformity constraint is represented here as PU/+ɾ/. Constructing a complex onset with [ɾ] in second position achieves this. The inserted consonant should in principle be the default consonant, and [d] fits

(6)

		CONTIG	MSD6	PU/+ɾ/	SYLCON	DEPV	DEPC
(a)	*folro* ˈfolR+o						
	ˈfo.lɾu		*!				
	☞ ˈfol.ru						
	ˈfo.lə̲.ɾu	*!				*	
	ˈfol.d̲ɾu	*!					*
(b)	*voldrem* ˈbɔl+R+ˈɛ+m						
	i. bul.ˈrɛm			*!			
	ii. bu.ˈlrɛm		*!				
	iii. ☞bul.ˈd̲ɾɛm						*
	iv. bu.lə̲.ˈrɛm					*!	
(c)	*Enric* enˈRik						
	ən.ˈd̲ɾik	*!					*
	ə.ˈnɾik		*!				
	☞ ən.ˈrik				*		

CONTIG, MSD6, PU/+ɾ/ » SYLCON, DEPV » DEPC

this description inasmuch as it is a denti-alveolar stop. A pure 'default' or unspecified consonant might be expected to be voiceless [t]. On the other hand, it is reasonable to suggest that [ltɾ], with two switches of voicing, is a more marked cluster than [ldɾ], with no voicing contour. Another perspective would be to see the epenthetic consonant not as a DEP violation but as an INTEGRITY violation. Thus /$l_1ɾ_2$/ would be realized as [$l_1d_{1,2}ɾ_2$], with [d] the result of the features common to [l] and [ɾ] (denti-alveolar place, voice) together with those features that allow the whole cluster to meet the syllable and sonority constraints (viz. [−cont], [−son]). I do not explore these alternatives further here. The interaction of the constraints now mentioned can be seen in tableau (6). Only candidates with the correct rhotic for the syllabic context are considered. Example (6b) shows DEPV » DEPC. Epenthetic segments are distinguished by underlining.

Tableau (7) illustrates, with *pertanyerem* ‘we shall belong’ and *vencerem* ‘we shall overcome’, the pattern of constraints leading to epenthesis of an unmarked vowel ([ə] in eastern Catalan, [e] in western Catalan).

(7)

		MSD6	*HETERORG	IDENTPA	PU/+ɾ/	*CC]$_\sigma$C	SYLCON	DEPV	DEPC
(a)	*pertanyerem* peR'taɲ+R+ɛ+m								
	i. pərtəɲ.'rɛm						*!		
	ii. pərtəɲ.'rɛm				*!		*		
	iii. pərtə.'ɲrɛm	*!							
	iv. pərtəɲ.'d̠rɛm		*!						*
	v. pərtən.'d̠rɛm			*!					*
	vi. ☞ pərtə.ɲə̠.'rɛm							*	
(b)	*vencerem* 'bɛns+R+ɛ+m								
	i. bəns.'rɛm					*!	*		
	ii. bəns.'rɛm				*!	*	*		
	iii. bən.'srɛm	*!							
	iv. bəns.'t̠rɛm					*!			*
	v. ☞ bən.sə̠.'rɛm							*	

CONTIG, **MSD6, *HETERORG, IDENTPA, PU/+ɾ/ » *CC]$_\sigma$C, SYLCON » DEPV** » DEPC

In (7) *HETERORG stands for the relevant coda place assimilation constraints (§6.2). In the *pertanyerem* case, epenthesis of [d] results in a *HETERORG violation (7a.iv), or, if this is avoided (7a.v), in fatal unfaithfulness to input place features. The comparison of (7a.i) and (7a.vi) shows SYLCON » DEPV. In *vencerem* (7b), whose root ends in an alveolar /s/, it is not *HETERORG that acts to rule out consonant epenthesis, but rather the constraint that penalizes internal two-consonant clusters: *CC]$_\sigma$C (§7.5 (24)). In fact, it is only in the (6b) case involving epenthesis of [d] in *voldrem*, etc. that PU/+ɾ/ is clearly active. If one were to regard Catalan /l/ as inherently less sonorous than trilled [r], then candidate (6b.i) would be excluded by SYLCON.

8.3 CODAS AND EPENTHESIS

In principle, admissible coda clusters are those which respect the sonority hierarchy set out in (4), as in the following examples (8) selected from Lloret

(2002a: 219–20). Nasal+sibilant clusters are admissible, so Minimum Sonority Distance seems not to apply to codas.

(8) -rC: *ferm* [ˈfɛrm] 'firm', *forn* [ˈforn] 'oven', *amorf* [əˈmɔrf] 'amorphous', *guerx* [ˈgɛrʃ] 'warped', *amarg* [əˈmark] 'bitter'
-lC: *salm* [ˈsalm] 'psalm', *golf* [ˈgolf] 'gulf', *Elx* [ˈɛlʃ] (toponym), *calb* [ˈkalp] 'bald', *talc* [ˈtalk] 'talc'
-NC: *romanx* [ruˈmanʲʃ] 'Rumantsch', *triomf* [triˈomf] 'triumph', *exempt* [əgˈzemt] 'exempt'
-sC: *bosc* [ˈbɔsk] 'wood', *Casp* [ˈkasp] (toponym), *llest* [ˈʎest] 'clever'

There are two problem issues. First, [s]/[z] may appear 'outside' a stop: word-final examples are those like *fix* [ˈfiks] 'stable', *ortodox* [urtuˈðɔks] 'orthodox', *bíceps* [ˈbisəps] 'biceps'; word-internal examples are *abstemi* [əps.ˈtɛmi] 'teetotal', *extern* [əks.ˈtɛrn] 'external', *adstrat* [ə͡ts.ˈtrat] 'adstratum'. There are even -CsC codas such as *text* [ˈtekst] 'text', *mixt* [ˈmikst] 'mixed', where [s] intervenes in a stop sequence that could not itself form a coda (*[tekt]). Second, not all sequences of falling sonority are in fact acceptable in codas. A glide may only be followed by an obstruent, not by another sonorant. Thus *naip* [ˈnajp], *caut* [ˈkawt] 'careful', *vuit* [ˈbujt] 'empty', *dijous* [diˈʒɔws] 'Thursday', and so on, are found, but the handful of examples with glide+sonorant are borrowings or archaisms which have alternative current forms and/or pronunciations (details in Wheeler 1987; Lloret 2002a: 219–22). A form such as *escriure* /skRiw+r/ [əsˈkɾiw.rə] 'write.INF', with epenthesis, illustrates the productive pattern of glide+sonorant clusters. The case of 'external' s/z exemplifies the special status of sibilants in syllable structure that is also seen in onsets in English (though not in Catalan): *spring*, *small*, etc. I do not offer a formal solution to this issue here, though a provisional expedient might be to allow /s/ two places in the coda sonority hierarchy, both its 'proper' place as in (4), and a special place outside the least sonorous group (obstruent stops), with, for example, a −1 value. For the glide+C clusters I invoke a special minimum sonority distance constraint for coda glides as in (6). The motivation for such a constraint requires further investigation.

(9) MINIMUM SONORITY DISTANCE (Coda glides) (MSDCG): In a coda sequence G_1C_2 (where G = glide), the value of $G_1 - C_2 \geq 11$ (values as in (4)).

The examples in (10) illustrate the types of cluster that provoke right-edge 'epenthesis'. Note that 'epenthesis' precedes inflections so the plural of *alegre* [əˈleɣrə] /aˈlegr/ 'cheerful' is *alegres* [əˈleɣrəs].

(10) (a) geminates (-CC of equal sonority)
esquerre /(e)sˈkɛRR/ [əsˈkɛ.rə] 'left', *batlle* /ˈbaʎʎ/ [ˈbaʎ.ʎə] 'mayor', *perenne* /peˈRɛnn/ [pəˈrɛn.nə] 'perennial', *summe* /ˈsumm/ [ˈsum.mə] 'supreme', *dotze* /ˈdoTz/ [ˈdod͡.d͡zə] 'twelve', *imatge* /iˈmaTʒ/ [iˈmad͡ʲ.d͡ʒə] 'image'

(b) -CC of equal sonority
correcte [kuˈrɛk.tə] 'correct', *apte* [ˈap.tə] 'suitable', *alumne* [əˈlum.nə] 'pupil', *folre* [ˈfol.rə] 'lining', *digne* /ˈdiɡn/ [diŋ.nə] 'worthy'

(c) -CC of rising sonority: onset
miracle [miˈɾa.klə] 'miracle', *obre* [ɔ.ˈβɾə] 'open.3SG.PR.IND', *monstre* [ˈmɔns.tɾə] 'monster', *amable* [əˈmab.blə] 'kind', *angle* [ˈan.glə] 'angle', *setembre* [səˈtem.bɾə] 'September', *múltiple* [ˈmulti.plə] 'multiple', *perdre* /ˈpɛRd+ɾ/ [ˈpɛɾ.ðɾə] 'lose.INF'

(d) -CC of rising sonority: coda+onset
-Tz, -Tʒ: see (10a); Kn: see (10b); *ritme* /ˈRiTm/ [ˈrid.mə] ∼ [ˈrim.mə] 'rhythm', *feminisme* /+ˈizm/ [fəmiˈniz.mə] 'feminism'

(e) -CC of falling sonority: glide coda+sonorant onset
aire /ˈajR/ [ˈaj.ɾə] 'air', *lliure* /ˈʎiwR/ [ˈʎiw.ɾə] 'free', *faune* /ˈfawn/ [ˈfaw.nə] 'faun', *guilleume* [giˈʎɛw.mə] 'rabbet plane', *retaule* [rəˈtaw.lə] 'altarpiece'

In addition to the cases in (10) there is some evidence that clusters of sonorant+voiced sibilant provoke epenthesis (Wheeler 1975/1979: 17–18; Lloret 2002a: 230–232, who includes further voiced clusters, as in /zb/ *bisbe* 'bishop' and /wd/ *laude* 'finding'). The evidence is most consistent in the case of the alveolo-palatal sibilant /ʒ/ (11a; no counter-examples). After consonant+alveolar /z/, alongside 12 cases of epenthesis (11b) there are three counter-examples (11c), for each of which an 'excuse' is available.

(11) (a) -C.ʒ falling sonority: coda+onset
diumenge [diwˈmenʲ.ʒə] 'Sunday', *conserge* [kunˈsɛr.ʒə], *greuge* [ˈɡɾew.ʒə] 'complaint' (total 16 roots: no counter-examples[3])

(b) -C.z falling sonority: coda+onset
colze [ˈkol.zə] 'elbow' (and four other /-lz/ roots), *catorze* [kəˈtor.zə] 'fourteen', *bronze* [ˈbɾon.zə] 'bronze' (and five other /-nz/ roots) (total 12 roots)

(c) non-epenthesis after -Cz
fons /ˈfonz/ [ˈfons] 'bottom' (cf. *enfonsar* [əmfunˈza] 'to sink'), *dins* /ˈdinz/ [ˈdins] 'within' (cf. *endinsar* [əndinˈza] 'to insert'), *brunz* [bɾuns] 'buzz.3SG.PR.IND'

Since clusters involving voiceless sibilants do not provoke epenthesis (see (8) above, and e.g. *temps* [ˈtems] 'weather', *llenç* [ˈʎens] 'linen'), epenthesis in the (11) examples could only follow directly from a version of MSDCG if the sonority hierarchy distinguished voiced from voiceless sibilants, and in such a way that sonorant+voiceless sibilant clusters passed the MSD constraint, while sonorant+voiced sibilant clusters did not. Yet it seems likely that other

[3] Counts derived from Mascaró & Rafel (1990).

considerations are involved in such cases. The clusters with epenthesis, in which the sibilant occurs in an onset, can contrast in voicing in a way that word-final clusters cannot. However, widespread coda voice neutralization suggests that Catalan grammar elsewhere attaches little importance to preserving voice in stem-final consonants. Another issue that arises with the (11a) type is how the cluster would be realized in the absence of epenthesis. The voiced alveolo-palatal /ʒ/ is realized as an affricate [d͡ʒ] ∼ [t͡ʃ] in word-final position, as in *boig* /bɔʒ/ [ˈbɔt͡ʃ] 'mad.M' ∼ *boja* [ˈbɔʒə] 'mad.F'. If this principle of distribution were applied to sonorant+fricative clusters as in *diumenge* 'Sunday' the citation form might be *[diwˈmen t͡ʃ], which would violate the *THREE MANNER CODA constraint (§7.3 (13)). I do not attempt a more conclusive account of this anomaly here.[4]

Tableau (12) illustrates the general pattern of sonority-driven word-final vowel epenthesis in types (10a) and (10b), with the examples *batlle* 'mayor' and *apte* 'suitable'. Epenthetic Vowels are underlined.

(12)

	batlle ˈbaʎʎ	SONSEQ	CONTIG	MAX/UNIF	MSD6	SYLCON	DEPV
(a)	ˈbaʎʎ	*!					
	ˈbaʎ			*!			
	☞ ˈbaʎ.ʎə̲						*
	ˈbaʎə̲ʎ		*!				
(b)	*apte* ˈapt						
	ˈapt	*!					
	ˈap			*!			
	☞ ˈaptə̲						*
	ˈapə̲t		*!				

SONSEQ, CONTIG, MAX/UNIF, MSD6 » SYLCON » **DEPV**

Tableau (13) illustrates the clusters in (10c) and (10d) with the examples *miracle* 'miracle' and *ritme* 'rhythm'; candidates violating CONTIG and MAX or UNIFORMITY are omitted.

[4] The historical reason for 'epenthesis' in (11a) and (11b), in many cases, is that originally each of the consonants in the cluster was in onset position, in a ˈVC$_1$VC$_2$V sequence, to which either syncope (> ˈVC$_1$C$_2$V) or apocope (> ˈVC$_1$VC$_2$), but not both, applied; as in ˈMONACHUM > *monge* 'monk'.

(13)

		SonSeq	MSD6	SylCon	DepV
(a)	*miracle* mi'ɾakl				
	mi'ɾakl	*			
	mi'ɾak.lə̯			*!	*
	☞ mi'ɾa.klə̯				*
(b)	*ritme* 'ridm				
	'ridm	*!			
	☞ 'rid.mə̯			*	*
	'ri.dmə̯		*!		*

SonSeq, Contig, Max/Unif, **MSD6 » SylCon » DepV**

Tableau (14) introduces the constraint ContigM(orph); it selects between the last two candidates, neither of which violates ContigI-O. ContigM penalizes candidates where morphemes that are supposed to be adjacent get separated. It will be relevant in all the cases of final epenthesis in infinitives. It may be interpreted in the present context as an abbreviation for a morphological alignment constraint Align(Suffix, L, Stem, R). I examine ContigM(orph) again in §11.4.2, where its position above DepV is demonstrated.

(14)

perdre 'pɛRd+ɾ	SonSeq	MSD6	SylCon	ContigM	DepV
'pɛɾðɾ	*!				
'pɛrd.ɾə̯			*!		*
☞ 'pɛɾ.ðɾə̯					*
'pɛɾ.ðə̯ɾ				*!	*

SonSeq, Contig, Max/Unif, MSD6 » **SylCon** »
ContigM » DepV

An infinitive form like *vèncer* 'overcome.inf' /ˈbɛns+ɾ/ [ˈbɛnsə] offers a good example of surface opacity, inasmuch as the elements of the sequence that would justify epenthesis, namely, /-ns+ɾ#/, are not all present in the surface form. As explained in §10.2.2, in most varieties of Catalan the /+ɾ/ suffix in the infinitive is pronounced only before a pronominal enclitic. I suggest in §10.2.2 that there is an anti-alignment constraint disfavouring infinitive /+ɾ/ at the right edge of a prosodic word (*AlignR/+ɾ/). This constraint 'competes' with RealizeMorph ((24b) in §11.4.2) to the effect that [ˈbɛnsə] with very unfaithful realization of /+ɾ/ is better than any of the alternative candidates. This state of affairs is illustrated in tableau (15).

(15)

	vèncer ˈbɛns+ɾ₁	SonSeq	MSD6	RealM	*AlignR/+ɾ/	SylCon	ContigM	Integrity	DepV	DepC
(a)	ˈbɛnsɾ	*!			*					
(b)	ˈbɛn.sɾə̲		*!						*	
(c)	ˈbɛns.rə̲					*!			*	
(d)	ˈbɛns.t̲ɾə̲						*!		*	*
(e)	ˈbɛn.sə̲₁ɾ₁				*!			*		
(f)	ˈbɛns			*!						
(g)	ˈbɛn.sə̲₁ɾ₁ə̲							*	*!	
(h)	☞ ˈbɛn.sə̲₁							*		

SonSeq, MSD6, RealM, *AlignR/+ɾ/, **SylCon** »
ContigM » **Integrity**, **DepV** » **DepC**

In (15) Integrity is introduced alongside DepV, to acknowledge the suggested interpretation of epenthetic [ə] as arising from 'splitting'. It is this interpretation which is necessary if [ə] in [ˈbɛnsə] is to count as 'realizing' the morpheme /+ɾ/, whereby candidate (15h) does not incur a violation of RealizeM. There is more than one approach to extending this concept to the regular realization of infinitives in Catalan, without [ɾ] in the majority of their realizations, as in *parlar* [pərˈla] 'to speak', where we might want to say that the infinitive is 'realized' in the element [ˈa] by which this form is distinguished from others in the paradigm, in particular from the root /parl/. I do not pursue further here this question that would take us deeper into the realm of morphological analysis.

Tableau (16) illustrates examples of the type in (10e), with *retaule* 'altarpiece' and *escriure* 'write.INF'. The top candidates are eliminated not by SonSeq, but by Minimum Sonority Distance (Coda glides) (MSDCG) (9). I retain here the conventional interpretation of 'epenthesis' as involving Dep violation. With a 'splitting' interpretation, the last candidate in (16b) would lose on *AlignR/+ɾ/ rather than on ContigM.

(16)

(a) *retaule* Reˈtawl	SonSeq	Contig	MSDCG	MSD6	SylCon	ContigM	DepV
rəˈtawl			*!				
☞ rəˈtaw.lə							*
rəˈta.wlə	*!			*			*
rəˈta.wəl		*!					*

(b)

escriure ˈskRiw+ɾ	SonSeq	Contig	MSDCG	MSD6	SylCon	ContigM	DepV
əsˈkɾiwɾ			*!				
☞ əsˈkɾiw.ɾə							*
əsˈkɾi.wɾə	*!			*			*
əsˈkɾi.wəɾ						*!	*

SonSeq, Contig, Max/Unif, **MSDGC,** MSD6 » SylCon » **ContigM** » **DepV**

Observe that, by Richness of the Base the input forms for all the monomorphemic examples in this section could equally well contain a final default vowel. For input /ˈapte/ in (12b), for example, the same output form ([ˈaptə]) would win, but would display no violation of DepV. But before a vowel-initial affix (as in /ˈapte+a/ 'suitable.F') such an input form would need 'repair' by 'truncation', through such constraints as penalize hiatus between non-high unrounded unstressed vowels. The fact is that /apt/ and /apte/ are not in contrast, so it is immaterial which is taken as 'underlying'.

8.4 SIBILANTS AND EPENTHESIS

The grammar of Catalan adopts a number of strategies to avoid word-internal clusters involving more than one sibilant. For example, the adjective *feliç* [fəˈlis] 'happy', which does not mark gender in the singular, has plural forms *feliços* 'happy.M.pl', *felices* 'happy.F.pl' that make use of the gender markers /o/ and /a/ ([ə]) to avoid */felis+s/, which would be pronounced [fəˈlis] like the singular, [ss] being barred by *GemSib. As mentioned in Chapter 7, in many varieties the plural of *gust* 'taste' is *gustos*, rather than *gusts*. These varieties use the /o/ masculine gender allomorph to avoid /sCs/ clusters. Sometimes fusion/reduction of adjacent sibilants (or sibilant clusters) is not avoided: *goig* [ˈgɔt͡ʃ] 'song to the Virgin' has plural *goigs* [ˈgɔt͡ʃ]; similarly *falç* [ˈfals] 'sickle' has plural *falçs* [ˈfals], and *dilluns* [diˈʎuns] 'Monday' has plural *dilluns* [diˈʎuns] (Wheeler 1975/9: 22–7). *Aquests* 'these.M' is generally pronounced [əˈkɛt͡s]. Within nominals, sibilant cluster avoidance is achieved, when it is achieved, by the use of morphological alternatives with syllable nuclei. There is much of morphological interest in these patterns, which are discussed along with other morphophonological patterns in nominals by Lloret & Viaplana (1992) and Lloret (1998). Though phonological conditions are involved, the main issues are morphological, and for this reason I do not discuss them further here. In verbs, however, sibilant clusters are consistently split up by vowel epenthesis, wherever such a cluster would arise for morphological reasons. The cases involve ten

sibilant-final stems (17a) which are followed immediately by the 2SG.PR.IND suffix /+z/ in the athematic conjugations, along with all the regular verbs of Conjugation IIIa, which contain the stem-extending affix /+ɛʃ/, as in the example in (17b).

(17) (a) *apareixes* /apaˈɾɛʃ+z/ [əpəˈɾɛʃəs] 'appear.2SG.PR.IND'
coneixes /koˈnɛʃ+z/ [kuˈnɛʃəs] 'know.2SG.PR.IND'
creixes /ˈkɾeʃ+z/ [ˈkɾeʃəs] 'grow.2SG.PR.IND'
cuses /ˈkuz+z/ [ˈkuzəs] 'sew.2SG.PR.IND'
fuges /ˈfuʒ+z/ [ˈfuʒəs] 'flee.2SG.PR.IND'
mereixes /meˈɾɛʃ+z/ [məˈɾɛʃəs] 'deserve.2SG.PR.IND'
neixes /ˈneʃ+z/ [ˈneʃəs] 'be-born.2SG.PR.IND'
torces /ˈtɔrs+z/ [ˈtɔrsəs] 'twist.2SG.PR.IND'
tusses /ˈtus+z/ [ˈtusəs] 'cough.2SG.PR.IND'
vences /ˈbɛns+z/ [ˈbɛnsəs] 'defeat.2SG.PR.IND'
(b) *prefereixes* /pɾefeˈɾ+ɛʃ+z/ [pɾəfə ˈɾɛʃəs] 'prefer.2SG.PR.IND'

The constraint hierarchy that prefers epenthesis here is relatively straightforward (18).

(18)

creixes ˈkɾeʃ$_1$+z$_2$	*{ʃ}{ʃKPT}	*GEMSib	REALM	CONTIGM	DEPV
ˈkɾeʃs	*!				
ˈkɾeʃ			*!		
ˈkɾess		*!			
☞ ˈkɾeʃəs				*	*

***{ʃ}{ʃKPT}, *GEMSib, REALM » CONTIGM » DEPV**

In (18) *{ʃ}{ʃKPT} is the heterorganic cluster constraint (11) from §6.2, *GEMSib is the constraint mentioned in §6.5.1 barring geminate sibilant fricatives, and REALM stands for the constraint or constraints that preserve the distinctiveness of each morpheme (taken up in §11.4.2 (24b)). Candidates with fusion such as [ˈkɾeʃ$_{12}$] or [ˈkɾes$_{12}$], or with elision of the stem-final consonant [ˈkɾes$_2$], are taken to be excluded by UNIFORMITY constraints, penalizing fusion, or PARADIGM UNIFORMITY constraints, penalizing allomorphy, whether of the stem or, more particularly, of the affix /+z/, whose place features are never compromised.

8.5 MINIMUM SONORITY DISTANCE AND VOICED STOP GEMINATION

One of the less well understood phenomena of Catalan phonology is the preference for geminate voiced stops before [l]: *possible* [puˈsib.blə] 'possible', *regla* [ˈreg.glə], and so on. Gemination is characteristic of eastern Catalan varieties and much of north-western Catalan. By contrast, Valencian and south-western parts of Catalonia use simple non-strident voiced fricatives in the corresponding items: [poˈsi.βle], [ˈre.ɣla] (Pradilla Cardona 2002: 296–8). The geminate stop may be pronounced voiceless in Barcelona and areas to the north of it, though this pronunciation is stigmatized. In the same areas intervocalic /pl/ and /kl/ may also geminate, for example, in *miracle* [miˈɾak.klə] 'miracle', leading to neutralization of voicing contrast in stop-lateral clusters in those varieties. Gemination is found in stressed syllables (19a, col. 1) and in the same roots when they appear in pretonic position (19a, col. 2), but not in syllables which are always unstressed (19b). In syllables which are unstressed in all related forms, fricative onset clusters [.βl], [.ɣl] are found. A few words lack expected gemination in stressed syllables (19c) (Wheeler 1981: 611–16; Mascaró 1987b; Lloret 1992).

(19) (a)

doble [ˈdob.blə] 'double'	*doblar* [dubˈbla] 'double.INF', *doblegar* [dub.bləˈɣa] 'fold.INF'
regla [ˈreg.glə] 'rule'	*reglar* [rəgˈgla] 'regulate.INF', *arreglar* [ərəgˈgla] 'arrange.INF'
diable [diˈab.blə] 'devil'	*diableria* [diəb.bləˈɾiə] 'devilry', *diablessa* [diəbˈblesə] 'she-devil'
república [rəˈpub.blikə] 'republic'	*republicà* [rəpub.bliˈka] 'republican'

(b) *deglutir* [də.ɣluˈti] 'swallow.INF'
negligir [nə.ɣliˈʒi] 'neglect.INF'
jeroglífic [ʒəɾuˈɣlifik] 'hieroglyphic'
oblidar [u.βliˈða] 'forget.INF'
problema [pɾuˈβlɛmə] 'problem'

(c) *ègloga* [ˈɛ.ɣluɣə] 'eclogue' /ˈɛglog+a/
bíblia [ˈbi.βliə] 'bible' /ˈbibli+a/
Publi [ˈpu.βli] (given name) /ˈpubli/
sigla [ˈsi.ɣlə] 'initial' /ˈsigl+a/

With respect to the exceptions in (19c), analysts such as Mascaró (1987b) and Lloret (1992) have remarked that in the first three of them the lateral cluster is not root-final, whereas it is in 'regular' examples of the (19a) type. In derivational GP this restriction was built into the gemination rule. In OT it might be modelled via an alignment constraint, such as ALIGNR(Root, σ).[5] However, this interpretation

[5] This would require interpreting *regla* [ˈreg.gl|ə] as better aligned than *[ˈre.ɣl|ə], on the grounds that the former has one more correspondent element aligned with the right edge of a syllable than the

requires that *república* [rəˈpub.blikə] 'republic' be analysed as morphologically complex, with /-publ/ ending the root; that is, the structure would have to be /re+ˈpubl+ik+a/, with the root /ˈpubl/, presumably, being an allomorph of /ˈpɔbl/ *poble* 'people' such as might also be seen in *públic* [ˈpub.blik] 'public'. The independent motivation for such an analysis is not strong. As well as the 'under-application' exceptions in (19c), Recasens (1993: 158) remarks on 'over-application' as well, in *sublevar* [sub.bləˈβa] 'to revolt' (a Spanish borrowing absent from standard dictionaries, though the Spanish pronunciation has no geminate: [su.βle.ˈβaɾ]), and in non-standard *iglésia* [ig.ˈglezjə] 'church'. He suggests social factors and style levels may be needed to account for both types of 'irregularity'.

Confining my analysis to the regular cases of (19a) and (19b), some sense can be made of the distribution if different versions ('strengths') of the Minimum Sonority Distance principle (5) are considered. The principle derives from a markedness scale for onset clusters. The least marked MSD constraint in universal grammar would be one with the highest value for sonority distance, permitting, for example, only voiceless stops before vowels. The most marked version (MSD0) would require no sonority distance at all between onset segments. I claim that, while MSD6 is the MSD constraint that is generally active in Catalan, in certain contexts, a less marked constraint, MSD7 may emerge, which, following the scale in (4) would penalize [.βl] and [.ɣl] onsets. This constraint, I believe, interacts with other markedness constraints that favour or disfavour gemination, together with a derivational faithfulness constraint, together set out in (20). I mention also LAZY, the constraint favouring lenition of voiced non-strident obstruents that is the focus of §10.1, though it is not crucial here.

(20) (a) MSD7: In an onset sequence C_1C_2, the value of $C_2 - C_1 \geq 7$
(b) *GEMINATE (*GEM): No geminates.
(c) STRESS-TO-WEIGHT PRINCIPLE (SWP): Stressed syllables are heavy (contain more than one mora).
(d) UNSTRESSED-LIGHT PRINCIPLE (ULP): Unstressed syllables are light (contain no more than one mora).
(e) IDENTBase-Derivative (IDENTB-D): A form in a morphological derivation is identical with its corresponding base.

I suggest also a new correspondence constraint DISTINCT which reflects a situation also captured in some approaches by MAXμ, but which does not entail prior syllabification, and thus mora assignment.

(21) DISTINCT (DIST): Every element of S_1 has a distinct correspondent in S_2. (DISTINCT allows $a_1b_2 \mathcal{R} x_1y_{12}$ or $a_1b_2 \mathcal{R} x_{12}y_2$ or $a_1b_2 \mathcal{R} x_{12}y_{12}$, but not $a_1b_2 \mathcal{R} x_{12}$. UNIFORMITY rules out all of these.)

latter has. This interpretation is not self-evidently correct. For *bíblia*, *[ˈbib.bli|.ə] would be no better or worse aligned than [ˈbi.βli|.ə].

The conceptual content of DISTINCT is that coalescence+breaking, resulting in a surface string of the same number of segments, is less marked than pure coalescence, which involves segmental loss (or loss of a timing slot). For example, for input /a_1u_2/, DISTINCT favours [$ɔ_{12}ʊ_{12}$] (mutual assimilation) over [$ɔ_{12}$] or [o_{12}], or indeed over [a_{12}] or [u_{12}], which involve merger.

The STRESS-TO-WEIGHT PRINCIPLE (SWP) and the UNSTRESSED-LIGHT PRINCIPLE (ULP) are two of the set of four Weight/Stress markedness constraints that favour the coincidence of stress with heavy syllables and lack of stress with light syllables (see Kager 1999: 172, 268). The relationship between these constraints is illustrated in (22), where candidates for *doble* 'double', *doblega* 'fold.3SG.PR.IND', *oblida* 'forget.3SG.PR.IND', and *negre* [ˈnɛ.ɣɾə] /ˈnɛgR/ 'black' are evaluated.

(22)

		SONSEQ	LAZY	SYLCON	IDENTB-D	ULP	MSD7	*GEM	SWP
(a)	*doble* 'dobl								
	'doβl	*!							
	'do.βlə						*!		*
	'do.blə		*!						*
	'dob.lə		*!	*					
	☞'dob.blə							*	
(b)	*doblega* 'dobl+'ɛg+a								
	du.'blɛɣə		*!						
	du.'βlɛɣə				*!		*		
	☞dub.'b1ɛɣə					*		*	
	dub.'1ɛɣə			*!		*			
(c)	*oblida* o'blid+a								
	u.'bliðə		*!						
	ub.'liðə		*!	*					
	☞ u.'βliðə						*		
	ub.'bliðə					*!		*	
	u.'bliðə		*!						

(d) *negre* ˈnɛgR	SonSeq	Lazy	SylCon	IdentB-D	ULP	MSD7	*Gem	SWP
ˈnɛgɾ	*!	*						
ˈnɛ.gɾə		*!						*
ˈnɛɣ.ɾə			*!					
ˈnɛg.gɾə							*!	
☞ ˈnɛ.ɣɾə								*

SonSeq, Lazy, SylCon, IdentB-D » ULP » MSD7 » *Gem » SWP

The winning candidates in (22a) and (22b) violate *Gem, but the alternatives are worse. MSD7 fatally penalizes [ˈdo.βlə] in (22a). Thus, in those varieties where [ˈdo.βlə] is indeed the winner *Gem outranks MSD7. In (22b) the second candidate is eliminated since it deviates further than the winner from root identity, despite having no geminate and an unstressed syllable which is light. In (22c), however, IdentB-D is not involved, and the cluster with a simple fricative wins, despite MSD7 violation. The SWP does not in fact select any winners in (22).

To illustrate the role of Distinct in relation to the other constraints mentioned in this section, consider candidates for *pitja* /ˈpidʒ+a/ [ˈpid͡dʒə] 'press. 3sg.pr.ind' and *despatx* /desˈpatʃ/ [dəsˈpa t͡ʃ] 'office' in (23).

(23)

(a) *pitja* ˈpid$_1$ʒ$_2$+a	SonSeq	Lazy	SylCon	IdentB-D	DepV	ULP	MSD7	Distinct	*Gem	SWP	Unif
i. ˈpid$_1$.d$_1$ʒ$_2$ə							*!		*		
ii. ˈpid$_1$.ʒ$_2$ə			*!							*	
iii. ˈpi.d$_1$ʒ$_2$ə							*!			*	
iv. ˈpi.d͡ʒ$_{12}$ə								*!		*	*
v. ☞ ˈpid͡$_1$.d͡ʒ$_2$ə									*		*
(b) *despatx* desˈpat$_1$ʃ$_2$											
i. dəsˈpat$_1$ʃ$_2$	*!										
ii. ☞ dəsˈpat͡ʃ$_{12}$								*			*
iii. dəsˈpat͡$_1$.t͡ʃ$_2$ə					*!				*		*

SonSeq, Lazy, SylCon, IdentB-D, **DepV** » ULP » MSD7, **Distinct » *Gem** » SWP

The winner in (23a) shows that some constraint outranks *Gem, which was seen in (22) to outrank SWP. Candidates (23a.i–iii) are excluded by sonority

constraints, SYLCON and MSD7 (or MSD6). Candidate (23a.iv), like the winner, (23a.v), displays coalescence, but in addition violates DISTINCT, which is thereby shown to rank above *GEM. In (23b), however, it is the winner (23b.ii) that shows the DISTINCT violation. Though (23b.iii) is inherently well-formed (compare *reprotxe* [rəˈpɾɔt.t͡ʃə] ∼ [rəˈpɾɔ.t͡ʃə] 'reproach'), it is not a good output for /desˈpatʃ/ since there is an alternative candidate available without epenthesis.

8.6 BALEARIC VERB FORMS WITH CODA CLUSTERS VIOLATING THE SONORITY SEQUENCING PRINCIPLE

Balearic Catalan differs from mainland varieties in having, in nearly all verb classes, no inflexional affix for the first person singular present indicative: stems appear in their bare form, for example, *deman* /deman/ 'request.1SG.PR.IND' (where continental varieties have *demano* ∼ *demane* ∼ *demani*). Stems with final clusters violating the sonority sequencing principle, or MSDCG (9), might therefore be expected to display epenthesis. So, just as the noun root /ˈfiltR/ *filtre* 'filter' is pronounced [ˈfiltɾə], so we might expect the verb /ˈfiltR/ 'filter.1SG.PR.IND' to be pronounced *[ˈfiltɾə]; as /ˈʎiwR/ *lliure* 'free' is [ˈʎiwɾə], we would expect /ˈʎiwR/ 'deliver.1SG.PR.IND' to be pronounced *[ˈʎiwɾə]. We find instead *filtr* [ˈfiltr̥] and *lliur* [ˈʎiwɾ] as verb forms. In (24) are further examples, mostly drawn from Lloret (2004); as in *filtr* [ˈfiltr̥] utterance-final sonorants are voiceless after a voiceless stop (cf. §5.2).[6]

(24) Balearic 1SG present indicative forms violating sonority sequencing
  (a) geminates (-CC of equal sonority)
    corr [ˈkor] /ˈkoRR/ 'run'[7]
    design [dəˈzinː] 'design'
    inject [inʲˈʒetː] 'inject'
    parl [ˈpalː] 'speak'
    vetl [ˈvəlː] 'stay awake'
  (b) -CC of rising sonority: (onset in other forms of the paradigm)
    arregl [əˈrekl] 'arrange'
    camufl [cəˈmufl] 'disguise'
    entr [ˈəntɾ] 'enter'
    mescl [ˈməskl] 'blend'
    mostr [ˈmɔstɾ] 'show'
    obr [ˈɔpɾ] 'open'
    sembr [ˈsəmbɾ] ∼ [ˈsəmpɾ] 'sow'
  (c) -CC of rising sonority: (coda+onset in other forms)
    blasm [ˈblazm] 'blame'

[6] The account offered in this section owes much to Lloret (2004); I am also grateful to Maria-Rosa Lloret for fruitful discussion of the issues involved.

[7] In Balearic varieties word-final [r] corresponds only to /RR/; /-R/ is realized [ɾ] (if it is not deleted).

(d) -CC of falling sonority: (glide coda + sonorant onset in other forms)
enlair [ənˈlajɾ] 'raise'
enllaun [ənʲˈʎawn] 'can'
entaul [ənˈtawl] 'sit down to table'

In fact, it is not only in first person singular forms that epenthesis fails in verbs in Majorca.[8] There is a handful of verbs with cluster-final stems in conjugations II and III, where there is no vocalic affix in the second and third persons singular of the present indicative either. These forms also lack epenthesis (25).

(25) Majorcan present indicative singular forms

	córrer 'run'	*obrir* 'open'	*omplir* 'fill'
1SG	*corr* [ˈkor]	*obr* [ˈɔpɾ]	*umpl* [ˈumpl]
2SG	*corrs* [ˈkors]	*obrs* [ˈɔt͡s]	*umpls* [ˈuns]
3SG	*corr* [ˈkor]	*obr* [ˈɔpɾ]	*umpl* [ˈumpl]

On the other hand, infinitives of conjugation II, where the affix /+ɾ/ immediately follows the root, show epenthesis on the islands as they do on the mainland: *escriure* /ˈskRiw+ɾ/ [əsˈkɾiwɾə] 'write.INF', *rebre* /Rəb+ɾ/ [ˈrə.βɾə] 'receive.INF'. That is to say, to avoid epenthesis in Majorcan Catalan, a cluster must simply be root-final, in a verb. The forms in (25), by the way, show that the coda cluster constraints proposed for Majorcan in Chapter 7 need to be refined: while a triconsonantal coda like [mpl] can be preserved by Paradigm Uniformity in 1SG or 3SG forms, in verbs, the same cannot be said for *[bɾs] or *[(m)pls] in 2SG, which are strongly disfavoured. Ensuring the cluster is homorganic and voiceless –*[tɾ̥s], *[(n)tl̥s]– is not a sufficient improvement. Such clusters are, of course, exceptionally marked in terms of the sonority sequencing principle, since in [Vtɾ̥s] sonority falls from V to [t], then rises from [t] to [ɾ], then falls again from [ɾ] to [s]. The rhyme [Vmpl] at least has only one sonority reversal: sonority falls from V through [m] to [p], then rises again for [l]. It will be seen shortly that an undominated CODA SONORITY SWITCHBACK (*CSSw) constraint reflects this state of affairs.

The failure of epenthesis in Balearic verb forms is explained by Lloret (2002b; 2004) in terms of the Optimal Paradigms element of Correspondence Theory within Optimality (McCarthy 2005). This approach highlights fundamental differences between nominal and verbal inflection in Catalan (and in other Romance languages similarly). Nominals have a maximum of four inflectional forms (in the case of a lexeme which can distinguish gender): masculine and feminine singular and plural. Typically the masculine singular is suffixless, the feminine singular adds /+a/, and the plural adds /+z/ to the singular: thus, of the four members of the paradigm two (feminine) have a vowel-initial inflection, /+a/, /+a+z/, and two (masculine) have not, /+Ø/, /+Ø+z/. A verb, however,

[8] I am grateful to Clàudia Pons for reminding me that, in Minorca, epenthesis does occur in the 2SG and 3SG forms only of (25).

has forty-five inflectional forms (forty-two finite, in seven screeves or sub-paradigms of six person/number forms, plus three non-finite forms). Of these forty-five, a minimum of forty are vowel-initial, and in the default conjugation I, forty-four of them are. (Most conjugation II infinitives are consonant-initial—the suffix is /+ɾ/—and some 'strong' participles are consonant-initial.) A vowel-initial affix normally follows a consonant-final stem. If a stem were to end in unstressed /ə/ or /e/, anti-hiatus constraints would ensure that the final vowel disappeared before a vowel-initial suffix. So if the stem of the verb *filtrar* 'filter' were, say, /ˈfiltrə/, this particular form could only surface when no vowel-initial inflection followed, i.e. only in the first person singular present indicative. Elsewhere the surface allomorph would always be /filtɾ/. The Optimal Paradigms approach to inflection penalizes surface allomorphy (as do other aspects of Output–Output Correspondence theory). In particular, OP counts the number of inflectional forms each potential allomorph would appear in. Candidates for evaluation by OP intra-paradigmatic correspondence constraints consist of whole paradigms. The constraints evaluate the correspondence relation of the shared part of each paradigm member, i.e. the output stem, with that of every other paradigm member. The OP constraints that govern allomorphy where the difference between allomorphs is the presence or absence of segments are OP-MaxC (penalizing members of the paradigm with deleted consonants), OP-MaxV (penalizing members with deleted vowels), OP-DepC (penalizing members with inserted consonants), and OP-DepV (penalizing members with inserted vowels). Tableau (26), inspired by the tableaux of Lloret (2002b; 2004) evaluates just the present indicative of *filtrar* in Majorcan, assuming an input stem /filtɾ/. (I illustrate with just the present indicative because nothing in any of the other screeves would alter the evaluation here.) *Hiatus abbreviates high-ranked constraints penalizing hiatus. Stem portions standing in correspondence are given in bold; epenthetic elements are underlined. OP constraints score a violation for each pair of forms within a paradigm that differ in stem form.

(26)

	filtrar filtɾ.PR.IND	*Hiatus	OP-MaxC	OP-DepV	SonSeq	MSDCG	MaxC	DepV
(a)	☞ <**filtɾ**, **filtɾəs**, **filtɾə**, **filtɾam**, **filtɾaw**, **filtɾən**>				*			
(b)	<**filtɾ<u>ə</u>**, **filtɾəs**, **filtɾə**, **filtɾam**, **filtɾaw**, **filtɾən**>			5*!				*
(c)	<**filt**, **filtɾəs**, **filtɾə**, **filtɾam**, **filtɾaw**, **filtɾən**>		5*!				*	
(d)	<**filtɾ<u>ə</u>**, **filtɾ<u>ə</u>əs**, **filtɾ<u>ə</u>ə**, **filtɾ<u>ə</u>am**, **filtɾ<u>ə</u>aw**, **filtɾ<u>ə</u>ən**>	*** **!						*** ***

***Hiatus, OP-MaxC, OP-DepV** » **SonSeq**, MSDCG, MaxC, DepV

Paradigm candidates (26b) and (26c) conform to SonSeq, in the 1sg form, in (26b) by epenthesis, and in (26c) by deletion, but the fact that the OP constraints are ranked above SonSeq means that the paradigm with a SonSeq violation (26a) wins. Candidate paradigm (26d), which satisfies SonSeq as well as OP correspondence by optimizing epenthesis, is excluded from the competition by the highly ranked *Hiatus markedness constraint, which is violated by five of its six forms. In (27) I consider an alternative underlying stem representation /filtrə/. As [ə] and zero are not in contrast word-finally after non-coda clusters, I need to show that the same paradigm wins, whatever the input, given this constraint ranking (illustrated here with MaxV and DepC constraints, which are the ones that might in principle improve syllable structures, given /filtrə/ as input).

(27)

	filtrar filtɾə.PR.IND	*Hiatus	OP-MaxV	OP-DepC	SonSeq	MSDCG	MaxV	DepC
(a)	<**filtɾ**, **filtɾəs**, **filtɾə**, **filtɾam**, **filtɾaw**, **filtɾən**>				*!		*** ***	
(b)	<**filtɾə**, **filtɾəs**, **filtɾə**, **filtɾam**, **filtɾaw**, **filtɾən**>		5*!					
(c)	<**filtɾə**, **filtɾəəs**, **filtɾəə**, **filtɾəam**, **filtɾəaw**, **filtɾəən**>	*** **!						
(d)	<**filtɾə**, **filtɾəjəs**, **filtɾəjə**, **filtɾəjam**, **filtɾəjaw**, **filtɾəjən**>			5*!				*** **
(e)	☜ <**filtɾəj**, **filtɾəjəs**, **filtɾəjə**, **filtɾəjam**, **filtɾəjaw**, **filtɾəjən**>							*** ***

In fact, (27) shows that, given input /filtrə/ and only *Hiatus ranked alongside the OP constraints, the winner from an OP point of view would be (27e) with insertion of a consonant throughout,[9] which avoids hiatus, avoids SonSeq violation, and provides a paradigm without allomorphy, despite violation of DepC across the board. The inference is that DepC ranks above SonSeq, as in (28).

(28) (Majorcan) ***Hiatus,** OP-MaxC, OP-DepV, **OP-MaxV, OP-DepC**, **DepC** » **SonSeq,** MSDCG, MaxC, **MaxV**, DepV

The same constraint hierarchy (28) in principle favours epenthesis in nominals. Consider *filtre* 'filter' (plural *filtres*) in (29): (29a) assumes input /filtɾ/ and (29b) underlying /filtrə/.

[9] The choice of anti-hiatic consonant is immaterial here; I choose [j] because there is evidence for insertion of final glides in 1sg.pr.ind in Balearic after vowel-final stems, as in *crear* 'create' *crei* [ˈkrej] 1sg.pr.ind, *crees* [ˈkreəs] 2sg.pr.ind, etc. In such cases, of course, the stem-final vowel is stressed, unlike /ˈfiltRə/.

(29)

		*HIATUS	OP-MAXC	OP-DEPV	OP-MAXV	OP-DEPC	DEPC	SONSEQ	MAXC	DEPV	MAXV
(a)	*filtre* filtɾ, pl. *filtres*										
	i. <**filtɾ**, **filtɾ**s>							*!*			
	ii. <**filtɾ**, **fil**s>		*!			*		*	**		
	iii. <**fil**, **fil**s>								*!***		
	iv. <**filtɾ**, **filtɾ̪**s>			*!	*			*		*	
	v. ☞ <**filtɾ̪**, **filtɾ̪**s>									**	
(b)	*filtre* filtɾə, pl. *filtres*										
	i. <**filtɾ**, **filtɾ**s>							*!*			**
	ii. <**filtɾ**, **filtɾə**s>			*!	*			*			*
	iii. ☞ <**filtɾ̪**, **filtɾə**s>										

(Balearic) *HIATUS, OP-MAXC, OP-DEPV, OP-MAXV, OP-DEPC, DEPC » **SONSEQ**, MSDCG, **MAXC » DEPV**, MAXV

Candidate (29a.iii) has a uniform paradigm and its members violate no sonority or hiatus constraints. Thus its MAXC violations must be fatal; hence MAXC outranks DEPV, the constraint violated by the winner (29a.v). The competition between two allomorphy-free paradigms (29a.i), which is faithful but with sonority violations, and (29a.v), which is well-formed but has epenthesis, shows that SONSEQ outranks DEPV.

(30)

	neutre (M.SG), *neutra* (F.SG), *neutres* (M.PL), *neutres* (F.PL) /nɛwtɾ/	*HIATUS	MAXC	*CSSW	OP-MAXC	OP-DEPV	OP-MAXV	OP-DEPC	DEPC	SONSEQ	DEPV	MAXV
(a)	<**nɛwtɾ**, **nɛwtɾ**s, **nɛw.tɾə**, **nɛw.tɾə**s>			*!						**		
(b)	<**nɛwt**, **nɛwt͡s**, **nɛw.tɾə**, **nɛw.tɾə**s>		*!*		4*			4*				
(c)	<**nɛwt**, **nɛwt͡s**, **nɛw.tə**, **nɛw.tə**s>		*!* **									
(d)	<**nɛwtɾ̪**, **nɛwtɾ̪**s, **nɛw.tɾ̪ə**, **nɛw.tɾ̪ə**s>	*!*										
(e)	☞ <**nɛw.tɾ̪**, **nɛw.tɾ̪**s, **nɛw.tɾə**, **nɛw.tɾə**s>					4*	4*				**	

***HIATUS, MAXC, *CSSW** » OP-MAXC, **OP-DEPV,** OP-MAXV, OP-DEPC, DEPC » SONSEQ, MSDCG » **DEPV,** MAXV

Finally I consider a full nominal paradigm, illustrated by the adjective *neutre* 'neutral' (30), assuming morphological structure /nɛwtɾ/, /nɛwtɾ+a/, /nɛwtɾ+s/, /nɛwtɾ+a+s/, for M.SG, F.SG, M.PL, and F.PL, respectively.[10] The winning output here is (30e) with epenthesis in the masculine forms, and thus allomorphy. Paradigm (30d) with epenthesis and no allomorphy satisfies SONSEQ but has two fatal hiatuses. (30b) is no worse a paradigm than the winner, (30e), since allomorphy is equal in the two models; but (30b) involves deletion. That this paradigm is disfavoured shows MAXC outranks DEPV. Candidate (30c) achieves paradigm uniformity by deletion across the board, and it incurs no hiatus or SONSEQ violations. Therefore, MAXC, the constraint by which (30c) loses, must rank higher than OP-DEPV or OP-MAXV, the highest constraints that the winner violates. The faithful candidate (30a) violates no OP constraints. Its masculine forms, of course, violate SONSEQ, but it was shown in (25) that SONSEQ is dominated by OP constraints (otherwise the syllabically anomalous 1SG present indicative forms would never surface). I suggest the constraint that eliminates (30a) is the CODA SONORITY SWITCHBACK constraint (*CSSw), mentioned above (25), or an equivalent constraint penalizing codas like /tɾs/. The relevant constraint is clearly not just *CCC]$_\sigma$ or *3MANNERCODA, because the data in (25) show that mpl]$_\sigma$ can survive in verbs, but not bɾs]$_\sigma$, or mpls]$_\sigma$. Thus the paradigm (30a) is eliminated because, while morphologically uniform, one of its members *[nɛwtɾs] is phonologically impermissible. If it were 'repaired' as *[nɛwt͡ɾs] as in (30b), (30c), the same form would incur a fatal MAXC violation.

Paradigms with the alternative input /nɛwtɾə/ are considered in (31). The first three rows are the paradigms without allomorphy: (31a) as (30a) has a fatal coda in the M.PL form; (31b) has fatal *HIATUS violations; while (31c) avoids hiatus and allomorphy at the cost of MAXC violations throughout the paradigm.

(31)

	neutre (M.SG), *neutra* (F.SG), *neutres* (M.PL), *neutres* (F.PL) /nɛwtɾə/	*HIATUS	MAXC	*CSSW	OP-MAXC	OP-DEPV	OP-MAXV	OP-DEPC	DEPC	SONSEQ	MAXV	DEPV
(a)	<**nɛwtɾ**, **nɛwtɾs**, **nɛw.tɾə**, **nɛw.tɾəs**>			*!						**	**	
(b)	<**nɛwtɾə**, **nɛwtɾəs** **nɛw.tɾəə**, **nɛw.tɾəəs**>	*!*										
(c)	<**nɛwtɾət**, **nɛwtɾəts** **nɛw.tɾətə**, **nɛw.tɾətəs**>		*!* **									
(d)	<**nɛw.tɾə̰**, **nɛw.tɾə̰s** ☞ **nɛw.tɾə**, **nɛw.tɾəs**>					4*	4*					**

[10] The forms, as pronounced in eastern Catalan, do not *require* an interpretation with the presence of contrastive feminine gender morphs. This is, however, the default pattern for an adjective, so the account must at least accommodate this possible morphological interpretation.

The constraint ranking presented in (30) and (31), though illustrated here with Majorcan pronunciation, is valid for other varieties *mutatis mutandis*, with the difference that SonSeq ranks high enough (above OP constraints) to drive epenthesis in those few verbs (those of (25)) with non-coda stem-final clusters and zero or non-vocalic suffix in 2sg and 3sg present indicative, despite the resulting surface (output) stem allomorphy: [kor] ∼ [korə], [ɔβɾ] ∼ [ɔβɾə], [ompl] ∼ [omplə] as in (32).[11] In Valencian such allomorphy in these three lexemes is avoided via the paradigms in (33), where the stems have been restructured to be vowel-final. Incidentally, this restructuring in turn introduces a new anomaly: verb roots with stress on a non-final syllable. (Other vowel-final roots are monosyllabic, like /ˈfa/ 'do', /ˈdu/ 'bring'.)

(32) Central Catalan present indicative singular

	córrer 'run' /koRR/	*obrir* 'open' /ɔbR/	*omplir* 'fill' /ompl/
1sg	*corro* [ˈko.ru]	*obro* [ˈɔ.βɾu]	*omplo* [ˈom.plu]
2sg	*corres* [ˈko.rəs]	*obres* [ˈɔ.βɾəs]	*omples* [ˈom.pləs]
3sg	*corre* [ˈko.rə]	*obre* [ˈɔ.βɾə]	*omple* [ˈom.plə]

(33) Valencian present indicative singular

	córrer 'run' /ˈkoRRe/	*obrir* 'open' /ˈɔbRi/	*omplir* 'fill' /ˈompli/
1sg	*córrec* [ˈko.rek]	*òbric* [ˈɔ.βɾik]	*ómplic* [ˈom.plik]
2sg	*corres* [ˈko.res]	*obris* [ˈɔ.βɾis]	*omplis* [ˈom.plis]
3sg	*corre* [ˈko.re]	*obri* [ˈɔ.βɾi]	*ompli* [ˈom.pli]

Forms similar to the Valencian 2sg and 3sg forms of *obrir* are in fact predominant in the Balearics also; somewhat less so are 2sg and 3sg forms of *córrer* with vowel-final stem allomorphs. These differences remind us that there are many, and unpredictable, ways of resolving conflicts between paradigm constraints and syllable structure constraints.

[11] Maria-Rosa Lloret (p.c.) draws my attention to the point that 'surface stem allomorphy' entails a particular analysis of the morphological structure of these forms. Other possibilities are (a) 'surface inflectional allomorphy' [s] ∼ [əs] for the 2sg suffix, [ə] ∼ Ø for the 3sg suffix; or (b) the presence of an 'empty morph' [ə] in the 2sg and 3sg forms. 'Empty morph' analyses seem to me inherently implausible psychologically. The choice to prefer a stem-allomorphy interpretation over an inflectional allomorphy interpretation reflects the fact that stem allomorphy is a characteristic phenomenon of irregular verb inflection in Catalan, while allomorphy of person–number inflection is extremely limited (Wheeler 2002).

9

STRESS, FEET, AND PHRASES

This chapter deals with the suprasegmental phenomena of stress and rhythm in Catalan. The main part offers an analysis of lexical accent; then in (§9.6) a provisional account of rhythm in the intonation phrase is suggested.

9.1 WORD STRESS: INTRODUCTION

With some limitations to be described shortly, the position of word stress in Catalan is contrastive. Thus *càrregues* [ˈkareɣes] 'load$_{N}$.PL', *carregues* [kaˈreɣes] 'load$_{V}$.2SG.PRS.IND', and *carregués* [kareˈɣes] 'load$_{V}$.1SG.PST.SUBJ' have the same vowels and consonants but differ in which syllable carries primary stress.[1] Each prosodic word (PrWd) has one stress, which falls on the rightmost stressable element, and every phonological word (= lexical word, except for items specified as clitics; see §9.2 below) is a prosodic word. Compounds, which consist of more than one lexical word, are stressed as sequences of prosodic words, i.e. as phrases, with each lexical word retaining stress, the rightmost element being most prominent. Lexical words may be free forms (bases that may stand alone or in combination with inflectional affixes), or combinations of stems with derivational affixes, such as suffixes, which carry their own lexical stress, or prefixes, which are unstressed. Some lexical words are bound stems that may only combine with other stems to form compounds, or with affixes to form complex words (§9.3).

The limitations on the position of word stress in Catalan words are of two kinds. First, the position of stress in all verb forms is determined by the inflectional morphology; stress falls on the last stressable affix—essentially exponents of tense or aspect categories, or of non-finiteness—or on the last syllable of the stem if there is no stressable affix. Secondly, prosodic words of all kinds are subject to a 'window' constraint: stress is never earlier in the word than antepenultimate (i.e. there is a three-syllable stress window), and if the penultimate syllable is heavy, stress cannot be antepenultimate. The analysis of the trisyllabic window and weight limitation offered here (§9.5) involves the prosodic category

[1] These example words are given in NW Catalan pronunciation, where unstressed /a/ and /e/ are not affected by vowel reduction. The corresponding forms in eastern Catalan are [ˈkarəɣəs], [kəˈrɛɣəs], and [kərəˈɣes].

'colon', situated between the levels of Prosodic Word and Foot, which groups feet in pairs (as argued by Green 1997).

It is usually assumed (e.g. by Serra 1996b; Bonet & Lloret 1998: 204) that in non-verb words penultimate stress is the unmarked pattern in vowel-final words (such as *rosa* [ˈrɔzə] 'rose' or *taxi* [ˈtaksi] 'taxi'), and final stress is unmarked in consonant-final words, such as *rebuig* [rəˈβut͡ʃ] 'rejection'. This 'unmarked' pattern is broadly in line with the orthography which indicates the position of stress in words displaying the marked patterns, such as *càrrega* [ˈkarəɣə] 'load', *violí* [biuˈli] 'violin', *fàcil* [ˈfasil] 'easy'. Discussions of stress in Catalan phonology have also generally assumed (rather than argued—see e.g. Serra 1996b; Bonet & Lloret 1998) that the basic pattern of the Catalan stress foot is trochaic. Oliva (1992), however, takes the view that the basic pattern is iambic, and I argue here that this view is correct, though the case for it is subtle rather than evident to inspection.

Non-lexical secondary stress is a central element of the rhythmic organization of phonological phrases, but does not directly form part of the prosodic organization of individual words. As Oliva (1992) points out, *finalitat* [finəliˈtat] 'purpose', which has word stress on the last syllable, may receive a secondary stress on either the first syllable ([ˌfinəliˈtat]) or the second ([fiˌnəliˈtat]), depending on the rhythmic context. Rhythmic secondary stress is irrelevant to the process of vowel reduction (§2.3) which affects all syllables lacking PrWd stress. Thus the second syllable in [fiˌnəliˈtat] undergoes vowel reduction to [ə] in eastern Catalan because it is not the syllable carrying the Prosodic Word stress; but this is no barrier to it bearing a rhythmic secondary stress.

9.2 UNSTRESSED WORDS

A small number of function words have no word stress: they are phonological clitics. It is apparently not possible to provide a watertight syntactic definition to delimit the class of phonological clitics in Catalan, though one can get close with a combination of phonological and syntactic criteria. The phonological criterion is that clitics are monosyllabic. In some cases it is hard to determine whether a function word has word stress or not, namely, when the vowel involved is /i/ or /u/—other vowels reveal lack of stress through vowel reduction in eastern varieties—since function words typically have a low degree of intonational prominence, and some, for syntactic reasons, cannot occur at the end of a phonological phrase where stress is always apparent. The complementizer *si* 'if, whether' and the conjunction *ni* 'neither, nor' are usually described as unstressed, as is the indefinite article *un*, but the latter has disyllabic forms such as *una* (F.SG), *unes* (F.PL) which always have more prominence on the first than the second syllable. (If *una*, *unes* were truly lexically unstressed, rhythmic stress might be expected to be able to fall on either syllable depending on the syntactic

context, as occurs in the case of clitic sequences such as *us en* (2PL.object+ partitive).) Absence of vowel reduction implies that the conjunction *o* [ɔ] 'or', the negative particle *no* [no] 'not', and the equative pivot *com* [kom] 'as' are lexically stressed; however, susceptibility to glide formation in the case of the conjunction *i* 'and', for example in *tu i ella* [tuˈjeʎə] 'you and she', betokens lack of word stress. Here the different stress behaviour of the coordinating conjunctions *o* 'or' and *i* 'and' confirms the unavailability of a purely syntactic definition of Catalan phonological clitics. Various phonological clitics are grouped syntactically in (1).

(1) Phonological clitics

(a) Conjunctions, complementizers	*i* 'and', ?*si* 'if', ?*ni* 'nor', *que* 'that (COMP)', 'that (RPRN)', (exclamative degree word, as in *que bé!* 'how good!')
(b) Verbal object pronouns	*lo* ∼ *el* (3SG.M), *la* (3SG.F), *li* (3SG.IO), *los* ∼ *els* (3PL.M/3PL.IO), *les* (3PL.F), *se* ∼ *es* (3.REFL), *me* ∼ *em* (1SG), *te* ∼ *et* (2SG), *nos* ∼ *ens* ∼ *mos* (1PL), *vos* ∼ *us* (2PL), *ne* ∼ *en* (partitive), *ho* (3N), *hi* (LOC/prepositional)
(c) Determiners	*lo* ∼ *el* (DEF.M.SG), *la* (DEF.F.SG), *los* ∼ *els* (DEF.M.PL), *les* (DEF.F.PL), *es* ∼ *so* (DEF.M.SG), *sa* (DEF.F.SG), *es* ∼ *ets* ∼ *sos* (DEF.M.PL), *ses* (DEF.F.PL), *en* (PERSONAL.M), *na* (PERSONAL.F), *mon* (1SG.POSS.M.SG), *ma* (1SG.POSS.F.SG), *mos* (1SG.POSS.M.PL), *mes* (1SG.POSS.F.PL), *ton* (2SG.POSS.M.SG), *ta* (2SG.POSS.F.SG), *tos* (2SG.POSS.M.PL), *tes* (2SG.POSS.F.PL), *son* (3.POSS.M.SG), *sa* (3.POSS.F.SG), *sos* (3.POSS.M.PL), *ses* (3.POSS.F.PL)
(d) Prepositions	*a* 'to', *amb* 'with', *de* 'of', *en* 'in', *per* 'for'
(e) Auxiliary verbs	The 2SG and 3SG present indicative of *ser* 'be' have non-syllabic post-vocalic enclitic forms, for example, *tu ets aquí* [ˈtud͡zəˈki] 'you are here', *no és veritat* [ˈnoz βəriˈtat] 'it isn't true'. In some varieties some or all of the present indicative forms of the perfect auxiliary *haver* are unstressed: [i] (1SG), *has* [əs] (2SG), *ha* [ə] (3SG), *hem* [əm] (1PL), *heu* [əw] (2PL), *han* [ən] (3PL).

As mentioned above, certain conjunctions and complementizers, namely, *o* [ˈɔ] 'or', *com* [ˈkom] 'as', which resemble syntactically the items of (1a), show word stress; for this reason, *si* and *ni*, whose high vowels do not unambiguously reveal 'unstressed' or 'low prominence', are marked with '?' in table (1). Among

pronouns, only verbal object pronouns (1b) are clitics.[2] Subject pronouns and prepositional object pronouns have word stress even if monosyllabic, for example, *jo* [ˈʒɔ] ~ [ˈjɔ] 'I', *ço* [ˈsɔ] (3SG.N), *amb vós* [əm ˈbos] 'with you', as have the interrogative and relative prepositional object pronouns *qui* [ˈki] 'who', 'whom', and *què* [ˈkɛ] 'what'. All the definite determiners (1c), which are also in fact monosyllabic, are clitics. Demonstrative determiners are generally at least disyllabic and stressed, such as *aquest* [əˈkɛst] 'this.M.SG', *aquelles* [əˈkɛʎəs] 'those.F.PL'; Valencian *est* 'this.M.SG' and *eix* 'that.M.SG' are usually pronounced disyllabically [ˈeste] and [ˈeʃe] respectively. Besides the five prepositions in (1d), which are monosyllabic and unstressed, only one preposition, *des* [ˈdes] 'since', is monosyllabic and stressed; all other prepositions, such as *sobre* [ˈsoβrə] 'on', *contra* [ˈkɔntrə] 'against', *darrere* [dəˈrerə] 'behind', are polysyllabic and stressed.

Catalan readily displays sequences of clitics, which can lead to quite lengthy strings of lexically unstressed syllables, for example, *que me les has amagades* [kə mə ləz əz əməˈɣaðəs] 'that you have hidden them from me'. Such sequences are subject to rhythmic stressing (see §9.6 below), so, depending on the context, such a sequence might be realized as follows (indicating rhythmic stress with [ˌ]): [ˌkə mə ˌləz əˌz əməˈɣaðəs] or [kə ˌmə ləz əˌz əməˈɣaðəs].

9.3 COMPOUNDS AND COMPLEX WORDS

9.3.1 *Compounds*

Compound words whose components are themselves prosodic words retain the stress location of their components. The main stress is that of the last stressed element, just as is the case in phrases. Thus, rhythmically, the compound *trencaclosques* [ˌtrɛŋkəˈklɔskəs] 'puzzle' does not differ from the verb phrase *trenca closques* [ˌtrɛŋkə ˈklɔskəs] 'breaks skulls'. The important differences between compounds and phrases are morphological, syntactic, and semantic, but not rhythmic. In (2) are displayed a few more examples of compounds, alongside their component elements.

(2) Examples of compound stress (free morphemes)

trescents [ˌtrɛˈsens] 'three hundred'[3]	*tres* 'three'+ *cents* 'hundreds'
home-llop [ˌɔməˈʎop] 'werewolf'	*home* 'man' + *llop* 'wolf'

[2] North Catalan and Balearic (Majorca/Minorca) enclitics bear primary stress; in some contexts, at least, the Balearic enclitics function rather as stressed inflections, involving loss of stress in stems, and vowel reduction, as in *mirar-la* 'see.inf-3SG.F.OBJ' [mirəˈlə] (central Catalan [miˈrarlə]). This quite complex system is the topic of a current doctoral research project by Pere Grimalt, at the Universitat de Barcelona.

[3] In Balearic varieties, the numbers 17 and 19 have the compound stress pattern: *desset* [ˌdɛˈsɛt] '17', *denou* [ˌdɛˈnɔw] '19', in which the first element is an allomorph of *deu* [ˈdɛw] '10'. In Valencian the numbers 17–19 are similarly compound, but exceptionally the initial element has primary stress:

copalta [ˌkoˈpaltə] ‘top hat’	*copa* ‘cup’ + *alta* ‘high’
malversar [ˌmalβərˈsa] ‘to embezzle’	*mal* ‘badly’ + *versar* ‘to turn’
entrepà [ˌentrəˈpa] ‘sandwich’ (also [əntrəˈpa] with *entre* treated as an affix rather than as a prosodic word;[4] Bonet & Lloret (1998: 40))	*entre* ‘between’ + *pa* ‘bread’
capficar [ˌkapfiˈka] ‘to insert’	*cap* ‘end’ + *ficar* ‘to put in’
migdia [ˌmid͡ʒˈðiə] ‘noon’	*mig* ‘middle’ + *dia* ‘day’

In addition to compounds of free morphemes such as those of (2), Catalan also has prosodic words which are not free morphemes, i.e. morphemes which bear word stress but which may combine only with other free or bound prosodic words. Many such elements have two variants (as indeed is the case with items like *entre* ‘between’ in (2)). One of these is stressed, like a prosodic word, while the other is unstressed, like an affix. Some examples of compounds involving bound elements are given in (3), arranged according to the bound or free nature of the elements. The focus is on the initial element; word stress is to be expected on the rightmost element whatever its bound or free status. Pronunciations are taken from Bruguera (1990), though considerable variation is reported in this area (Recasens 1993: 89).

(3) (a) Bound + free

Stressed first element

afroasiàtic [ˌafruəˈzjatik] ‘Afro-Asiatic’
angloamericà [ˌaŋgluəməriˈka] ‘Anglo-American’
arteriografia [ərˌtɛɾjuɣɾəˈfiə] ‘arteriography’
arteriosclerosi [ərˌtɛɾjusˈkləɾɔzi] ‘arteriosclerosis’
audiovisual [ˌawðjuβizuˈal] ‘audio-visual’
autosuficient [ˌawtusufiˈsjen] ‘self-sufficient’
cardiovascular [ˌkarðjuβəskuˈlar] ‘cardiovascular’
dendrocronologia [ˌdɛndɾukɾunuluˈʒiə] ‘dendrochronology’
grecoromà [ˌgɾekuruˈma] ‘Greco-Roman’
hagiotopònim [ˌaʒjutuˈpɔnim] ‘hagiotoponym’
heterosexual [ˌetəɾusəksuˈal] ‘heterosexual’

dèsset [ˈdɛˌsɛt] ‘17’, *dihuit* [ˈdiˌwit] ‘18’, *dènou* [ˈdɛˌnɔw]. Saragossà (2001) plausibly suggests that this stress pattern reflects rhythmic analogy with the trochaic pattern of the preceding members of the 11–19 paradigm (from *onze* ‘11’ through *setze* ‘16’).

[4] *Entrepà* [əntrəˈpa] is not an isolated alternative; Recasens (1993: 89) mentions variants in which *contra* ‘counter-’, *entre* ‘between, inter-’, *mal* ‘ill-’, *radio* ‘radio-’, *sobre* ‘over-’ are treated as unstressed affixes rather than free stressed stems in complex words. See Mascaró (1986a: 32) for observations on the morphological structure of these elements.

	macroeconomia [ˌmakɾuəkunuˈmiə] 'macro-economy'
	serbocroata [ˌsɛɾβukɾuˈatə] 'Serbo-Croat'
Unstressed first element	*anglosaxó* [əŋglusəkˈso] 'Anglo-Saxon'
	cardiografia [kərðjuɣɾəˈfiə] 'cardiography'
	gal·loromà [gəlluruˈma] 'Gallo-Roman'
	hagiografia [əʒjuɣɾəˈfiə] 'hagiography'
(b)	Bound + bound
Stressed first element	*autodidacte* [ˌawtuðiˈðaktə] 'self-taught'
	estereoscopi [əsˌtɛɾeusˈkɔpi] 'stereoscope'
	macrocèfal [ˌmakɾuˈsɛfəl] 'macrocephalic'
	semiluni [ˌsɛmiˈluni] 'half a lunar month'
Unstressed first element	*heterodox* [ətəɾuˈðɔks] 'heterodox'
	autonòmic [əwtuˈnɔmik] 'autonomic'
(c)	Free + bound
Stressed first element	*expressament* [əksˌpɾɛsəˈmen] 'purposely'
	principalment [pɾinsiˌpalˈmen] 'principally'

In type (3a), consisting of bound + free stems, maintenance of word stress on the initial element is probably the dominant pattern, with items variably shifting to manifest a single stress insofar as the morphological structure is disregarded. In type (3b), in which both elements are bound, single stress seems predominant, the bound status of both parts contributing to the sense that the whole is a complex rather than a compound word. But this has the consequence, itself leading to variation, that the same element, e.g. *hetero-*, is pronounced sometimes stressed and sometimes unstressed, in words of similar prosodic form; cf. *heterosexual* vs. *heterodox*. Type (3c) illustrates the only bound second element, *-ment*, that combines with a free first element. The pattern is extremely productive in deriving adverbs from adjectives. In fact, since the initial element in such compounds displays feminine singular gender marking, as well as word stress, there is no reason other than the orthography not to regard type (3c) merely as a noun phrase with a feminine nominal head which happens to be limited to this construction.

Compound, then, is not a category with any relevant prosodic consequences in Catalan. The necessary distinctions with prosodic consequences are between lexical stems and affixes and, among lexical stems, between free stems and bound stems. Free stems may stand alone or bear only inflectional affixes. Bound stems must combine with other stems or with derivational affixes: the combinatorial possibilities, or selection restrictions, of each bound stem are a matter for the lexicon. For example, bound stems such as *macro* 'macro-' or *vice* 'vice-' must be combined with a following element; while *scopi* '-scope' or *nom* '-nomous' need to combine with a preceding element.

9.3.2 *Complex words with affixes*

The stress pattern of complex words derived through affixation is straightforward. Each prosodic word has only one stress (by HEAD(PrWd)—see below), which is that of the rightmost stem or affix (by ENDRULE-R—see below). Stems and derivational affixes have the location of stress lexically marked, but the constraints HEAD(PrWd) and ENDRULE-R ensure that only one is realized on the surface. (There is thus no way that lexical stress on a true prefix could surface; see Mascaró (1986a: 33).) All derivational affixes are at least one syllable long, and nearly all of them are 'auto-stressed'. However, a few affixes, mostly deriving adjectives, such as *-i*, *-ic*, and *-id*, are 'prestressed',[5] i.e. they require stress on the syllable preceding them (stress which can in turn, of course, be overridden by the stress of a following suffix). Nominal inflectional affixes (gender and number) are unstressed. The relevant constraints are spelt out in (4).

(4) (a) HEAD(PrWd): Each prosodic word has a unique head (McCarthy 2002: 78).
(b) ENDRULE-R: The head foot is not followed by another foot within the prosodic word (McCarthy 2003b: 111).[6]
(c) FPH: Faithfulness to prosodic heads. Prosodic heads in the input are realized as such in the output (after Beckman 1998; Alderete 1999; Kager 2000).
(d) *LAPSE: *σ̆/__σ̆: Assign one violation mark for each pair of adjacent unstressed syllables (McCarthy 2003b: 106).

In the examples (5–7) below the position of word stress in inputs is indicated with underline. The footing of the candidates is in accordance with the constraints ALLFEETL and FOOTBIN (§9.5), though for the issues raised here it would not matter if it were otherwise; unlikely footings are not illustrated here. In the outputs central Catalan vowel reduction is applied to unstressed syllables.

Candidate (5a) retains the lexical prosodic heads of its stem and affix, in accord with FPH, but thereby violates HEAD(PrWd) which requires that there be only one head, and thus also ENDRULE-R, since the two heads cannot both be the rightmost foot. The comparison of candidates (5a) and (5b) demonstrates the ranking HEAD(PrWd) » FPH. The comparison of (5b) and (5c) illustrates

[5] The others (Mascaró 1986a: 24–31) are *-il* as in *erèctil* 'erectile', *-im* as in *cèntim* 'cent', *-it* as in *insòlit* 'unwonted', *-ul* as in *nòdul* 'nodule', and *-upl(e)* as in *quàdruple* 'quadruple'. There is one purely nominal prestressed suffix: *-ode* as in *elèctrode* 'electrode'.

[6] ENDRULE-R is offered by McCarthy as a categorical constraint to replace the gradient RIGHTMOST: ALIGN(HdFt, R, PrWd, R): 'The head foot is rightmost in the prosodic word' (Kager 1999: 167); *LAPSE constraints are also part of McCarthy's approach, after Kager (2002), to avoid gradient constraint evaluation.

(5) *italiana* 'Italian.F.SG' [itəliˈanə]:[7] /itali/ (cf. *Itàlia* 'Italy') /+an/ /+a/ (fem.)

	/itali]$_{St}$+an]$_{Af}$ +a]$_{In}$/	HEAD(PrWd)	FPH	ENDRULE-R	*LAPSE
(a)	iˈtaliˈanə (σˈσ)(σˈσ)σ	*!		*	
(b)	☞ itəliˈanə (σσ)(σˈσ)σ		*		**
(c)	itˈaliənə (σˈσ)(σσ)σ		*	*!	**
(d)	itəliənə (σσ)(σσ)σ	*!	**!		****
(e)	itəˈliənə (σ)(σˈσ)(σσ)		**!	*	***
(f)	itəliəˈnə (σ)(σσ)(σˈσ)		**!		***

the role of ENDRULE-R. Candidate (5d) violates HEAD(PrWd) by having no head at all. Candidates (5e) and (5f) ignore input stress location, fatally violating FPH twice; the head foot of (5e) is also misaligned. Candidate (5f) is well-aligned by ENDRULE-R, but its two violations of FPH show that it loses to candidate (5b).

Candidate (6a) again retains all the input stresses, but incurs multiple fatal violations of HEAD(PrWd). Candidate (6e) has no stress at all, also violating HEAD(PrWd). Candidates (6c) and (6d) fatally retain underlying stresses other than the rightmost, while the winning candidate, (6b), has one head, the rightmost foot in the word. Example (7), *sil·làbic* 'syllabic' illustrates a pre-stressed suffix applied to a stem with marked non-final stress. I ignore here the candidate with no stress.

Candidate (7a) again is ruled out by HEAD(PrWd) through having multiple heads. Candidate (7d) respects neither of the input stresses (two violations of FPH), though its head is rightmost. Candidates (7b) and (7c) tie on ENDRULE-R, since in each case the head foot is followed by another foot within the word; (7b) wins on *LAPSE.[8]

[7] Also [itəˈljanə] for younger speakers and/or more informal styles; see §3.2.4.7.

[8] Though I follow McCarthy (2003b) in replacing gradient evaluation by RIGHTMOST, the consequential account is perhaps not the most perspicuous way of expressing the fact that, among underlying stresses, it is the rightmost that surfaces.

(6) *mariner* [məri'ne] ‘sailor’: /maR/ (cf. *mar* ‘sea’) /+in/ (adjective-forming suffix) /+eR/ (agent suffix)

	/maR]$_{St}$ +in]$_{Af}$ +eR]$_{Af}$/	HEAD(PrWd)	FPH	ENDRULE-R	*LAPSE
(a)	'mə'ɾi'ne ('σ)('σ)('σ)	*!*		**	
(b)	☞ məɾi'ne (σ)(σ'σ)		**		*
(c)	mə'ɾine (σ'σ)(σ)		**	*!	*
(d)	'maɾinə ('σ)(σσ)		**	*!	*
(e)	məɾinə (σ)(σσ)	*!	***		**

(7) *sil·làbic* ‘syllabic’: /silab/ (cf. *sí l·laba* ‘syllable’) /+_ik/

	/silab]$_{St}$+_ik]$_{Af}$ /	HEAD(PrWd)	FPH	ENDRULE-R	*LAPSE
(a)	'si'la'βik ('σ)('σ)('σ)	*!*		**	
(b)	☞ si'laβik (σ'σ)(σ)		*	*	
(c)	'siləβik ('σ)(σσ)		*	*	*!
(d)	silə'βik (σ)(σ'σ)		**		*!

9.4 VERB STRESS

The position of stress in Catalan verbs is wholly predictable on the basis of their morphological structure, including conjugation class membership. Most tense, aspect, and mood suffixes, and non-finite suffixes (infinitive, gerund, participle) are auto-stressed: as with nominal derivational suffixes, main stress falls on the last of these. If no auto-stressed affix is present, stress falls on the final syllable of the stem (by HEAD(PrWd), ENDRULE-R (4), and IAMB—see (18) below). Non-syllabic affixes, such as the person/number endings /+z/ (2SG), /+m/ (1PL), /+w/ (2PL), /+n/ (3PL), the infinitive allomorph /+R/, and the participle allomorphs /+t/, /+st/, /+z/, unsurprisingly do not attract stress. The description of the verb-stress system depends to a degree on the analysis of the inflectional morphology, to which there are different reasonable approaches; see Perea (2002) and Wheeler (2002) for recent accounts of regular and irregular verbs respectively. Some

examples will illustrate these principles. Orthographic forms are used, as pronunciation of the vowels in particular varies dialectally in ways that are not relevant to the stress question.

Esperàvem, a paroxytone form, is the 1PL past imperfective indicative of *esperar* 'hope', a conjugation I verb. The stem is *esper-*; *-ave-* is one of two allomorphs of the conjugation I past imperfective indicative affix. (The other allomorph is *-ava-*; generally in Catalan inflectional morphology unstressed *-a-* and *-e-* alternate in a regular way, with *-e-* before a consonant, as here in *-àvem*, and with *-a* in word-final position.) The *-m* of *esperàvem* is the 1PL agreement marker. Stress regularly falls on the first vowel of the stressable affix.

Sentiria, a paroxytone form, is the 3SG conditional of *sentir* 'hear', a conjugation IIIb verb. The stem is *sent-*; *-iria* is a complex suffix containing *-ir-* which is also seen in the future tense (e.g. 3SG *sentirà*) and may be identical to the infinitive ending *-i*+R (*sentir*), followed by *-ia* (which may be identical to the *-ia* ∼ *-ie-* that marks past imperfective indicative in verbs from conjugations II and III). Third person singular agreement is not explicitly marked. Stress falls regularly on the first vowel (*-i-*) of the last suffix in the form.

Sabut, an oxytone form, is the participle (M.SG) of *saber* 'know', a conjugation II verb. The stem is *sab-*; the participle affix is /+ud/; stress falls again on the first vowel of the last affix.

The stress pattern of these and corresponding examples reflects the constraint ALIGNLSUFFIX (8).

(8) ALIGNLSUFFIX: ALIGN(DSuffix, L, PrWdHdμ, L). The left edge of a DSuffix is aligned with the head mora of a prosodic word.

In (8) the abbreviation DSuffix is used as a cover term for both the relevant verb suffixes and for nominal derivational suffixes, which obey the same constraint.[9] ALIGNLSUFFIX might also be formulated to say that the left edge of a DSuffix is aligned with a stressed syllable; that it is head of a prosodic word would follow from ENDRULE-R. The stress pattern of verbs manifests another alignment constraint (9) that imposes final stress on verb roots, dominating faithfulness in cases such as the root *carreg-* 'load' (see §9.1), which is stressed on the initial syllable in nominals.

(9) ALIGNRVSTEM: ALIGN(VStem, R, PrWdHd, R)

The stress pattern of present tense forms (indicative, subjunctive, and imperative) is a little more complicated. In these, the suffix is stressed only in the 1PL and 2PL forms. In the remainder, stress falls on the last syllable of the stem. The paradigms of the present indicative and subjunctive of *esperar* 'hope' are given

[9] Pre-stressed suffixes (here abbreviated PSuffix) reflect a left-right alignment constraint: ALIGN (PSuffix, L, PrWdHd, R).

below in standard central Catalan variants. (Dialect differences are not related to stress.)

(10)

	Indicative	Subjunctive
1SG	*es'pero*	*es'peri*
2SG	*es'peres*	*es'peris*
3SG	*es'pera*	*es'peri*
1PL	*espe'rem*	*espe'rem*
2PL	*espe'reu*	*espe'reu*
3PL	*es'peren*	*es'perin*

This irregular stress pattern in the present tense (which Catalan shares with most other Romance languages—see Roca 1999) does not reflect any natural prosodic or morphological tendency. It is a historical legacy from Latin, where the 1PL and 2PL endings were one syllable longer than the rest. In Latin this pattern was common to active verbal inflection in general; its reflex in Catalan has been restricted to the present tense. In the context of 1PL and 2PL the present tense markers *-a-*, *-e-*, *-i-* receive stress as DSuffixes in (8). In the other person/number forms of the present, any vowel affix (*-o*, *-e*, *-a*, *-i*) is unstressed just as nominal inflectional suffixes are. In fact, there is something to be gained by regarding the subjunctive suffix *-i-*, and its allomorphs *-iga* ∼ *-igue-*, *-ia* ∼ *-ie-*, as pre-stressed. The 2SG present subjunctive of *saber*, *'sàpigues*, displays an allomorph of the subjunctive affix which, being disyllabic, might be expected to receive stress on its leftmost element, as was seen above in the case of the *-ave-* element in *esperàvem*, for example. Unless another stressable suffix follows—as it does, for example, in *sapi'gueu* 'know.2PL.PRS.SUBJ' (allomorph of stem *sap-*, *-igu(e)-* PRS.SUBJ, *-e*-PRS (1PL/2PL), *-u* 2PL)—the subjunctive *-i-* is pre-stressed not only in the present tense but also wherever it occurs in the past subjunctive, namely, in the 2SG and throughout the plural—see paradigm corresponding to *esperar* 'hope' in (11). That is to say, one may see the unstressed nature of the *-i-* suffix in the present paradigm as reflecting not so much its 'present SG/3PL' character as its status as a pre-stressed suffix (PSuffix; see note 5).[10]

(11) Past subjunctive

espe'rés
espe'ressis
espe'rés
espe'réssim
espe'réssiu
espe'ressin

[10] This essentially pre-stressed character of subjunctive *-i-* is displayed in the otherwise anomalous stem stress in the 1PL and 2PL of the 'present tense' subjunctive of the perfect auxiliary (*hàgim* /aʒ+i+m/, *hàgiu* /aʒ+i+w/) and the past perfective auxiliary (*vàgim* /baʒ+i+m/, *vàgiu* /baʒ+i+w/).

Returning to the other unstressed verb suffixes, *-o-*, *-e-*, and *-a-* (which have a diverse range of person, number, and mood functions within the set of paradigms, and within different varieties), an approach might be to classify them as 'unstressed suffixes' (USuffixes) alongside the unstressed nominal suffixes of parallel form (such as *-o-*, *-e-*, masculine gender markers, *-a-* (*-e*-before *-s*): feminine gender marker). Unstressed suffixes would be subject to an anti-alignment constraint as in (12).

(12) *ALIGNFTUSUF: USuffixes are not aligned with the head of a prosodic word.

The non-syllabic suffixes such as the person/number endings /+z/ (2SG), /+m/ (1PL), /+w/ (2PL), /+n/ (3PL), the infinitive morpheme /+R/, and the participle allomorphs /+t/, /+st/, /+z/, and similarly the nominal plural /+z/, lacking a nucleus, would not attract stress anyway, though they may well find themselves in the rhyme of a syllable stressed for other reasons.

The present tense of verbs of conjugation IIIa is characterized by an affix (-ɛʃ- ~ -ɛsk- and other variants) whose role appears to be to bear word stress when no other stressable affix is present, thereby avoiding stress ever falling on the lexical stem.[11] (Conjugation IIIb consists of precisely those conjugation III verbs which do not display this element. Their stems are stressed in the relevant present SG/3PL forms.) The present indicative and subjunctive paradigms of a representative conjugation IIIa verb *referir* 'refer' is given in (13). Imperative forms are syncretistic with present indicative or subjunctive forms.

(13)

	Indicative	Subjunctive
1SG	*refeˈreixo*	*refeˈreixi*
2SG	*refeˈreixes*	*refeˈreixis*
3SG	*refeˈreix*	*refeˈreixi*
1PL	*refeˈrim*	*refeˈrim*
2PL	*refeˈriu*	*refeˈriu*
3PL	*refeˈreixen*	*refeˈreixin*

The general non-contrastiveness of stress location in verbs allows us to observe the interaction between stress and vowel epenthesis: epenthetic vowels are never stressed, despite resulting surface anomalies in stress patterns. The evidence comes in two places. One of them is in the verb *estar* 'stay'. The oxytone stress pattern of this verb in the singular and 3PL of the present indicative, 2SG *estàs* [əsˈtas], 3SG *està* [əsˈta], 3PL *estan* [əsˈtan], is consistent with the lexical stem being /st/ (or indeed /sta/—this interpretation may also be justified morphologically), rather than /est/. On the interpretation that the lexical stem is /st/ and that [ˈa] in these forms is a conjugation I allomorph of the present indicative marker, which is normally realized unstressed—for example in *jurar* 'swear': 2SG *ˈjures*,

[11] Because of vowel reduction, most verb stems which are stressed in some forms but not others display allomorphy, so a consistent pattern of stress on suffixes avoids stem allomorphy.

3SG ˈ*jura*, 3PL ˈ*juren*—the ungrammaticality of 2SG *[ˈestəs] reveals that the HEAD-DEPENDENCE constraint (14) dominates the *ALIGNFTUSUF constraint (12).

(14) HEAD-DEPENDENCE$_{IO}$ (HDDEP$_{IO}$): Every segment contained in a prosodic head in the output has a correspondent in the input (Alderete 1999: 36).
HDDEP$_{IO}$ » *ALIGNFTUSUF

It also follows that HEAD(PrWd) (4a) dominates *ALIGNFTUSUF, as shown in tableau (15). In (15) candidates for *estàs* are evaluated. The constraints penalizing syllable-initial /sC/ are abbreviated in (15) as SYLLABLESTRUCTURE (SYLSTR); the realization of unstressed vowels is that of eastern Catalan. The ranking HEAD(PrWd) » *ALIGNFTUSUF is likewise supported by present indicative forms of *dar* 'give' (2SG *das*, 3SG *da*, 3PL *dan*); this conjugation I verb also has an asyllabic stem /d/.

(15)

st+a+s	HEAD(PrWd)	HDDEP$_{IO}$	SYLSTR	*ALIGNFTUSUF
ˈstas			*!	*
stəs	*!		*	
☞ əsˈtas				*
ˈestəs		*!		
əstəs	*!			

The other type of data illustrating stress-epenthesis interaction involves conjugation II verb stems followed by the non-syllabic infinitive allomorph /+R/. In the case of verb stems whose final consonant may legitimately combine with /R/ to form an onset, an epenthetic vowel follows, as in paroxytone ˈ*rebre* 'receive.INF' (/Rɛb+R/) or ˈ*perdre* 'lose.INF' (/pɛRd+R/). Here too high-ranking HDDEP$_{IO}$ ensures that it is not the epenthetic vowel that receives stress (by HEAD(PrWd), ENDRULE-R (4), and IAMB—see (20) below). If the verb stem ends in a consonant that may not form an onset with /R/, epenthesis precedes /+R/, but the rightmost stressable element remains the last syllable of the stem, for example in *ésser* 'be' /es+R/, *néixer* 'be born' /neʃ+R/, *estrènyer* 'squeeze' /stRɛɲ+R/.

9.5 PROSODIC WORD CONSTRAINTS

As mentioned in §9.1, in Catalan the location of stress—one per prosodic word—in non-verbs is a lexical property. Not all conceivable stress locations are found, however, and the task of this section is to account for the gaps. One constraint has just been mentioned: *ALIGNFTUSUF: USuffixes are not aligned with the head of

a prosodic word. Inflectional affixes are, not unusually, 'extrametrical'. Where the base exists as an uninflected prosodic word this fact could be attributed to correspondence constraints, along the lines suggested by Alderete (2001: 191–3). For example, one might say that the stress profile of *cunyades* 'sisters-in-law' (/kuɲad/ /+e/ (F) /+z/ (PL)) is constrained to match that of the base *cunyat* [kuˈɲat] 'brother-in-law'. But this would not provide an account of the fact that the feminine gender marker /a/ ∼ /e/ is never stressed, even in words that have no correspondent without the agreement marker. *Fiˈgura* 'figure' is well-formed, but there is no form like **figuˈra* where *-a* is inflectional. Clearly, then, *ALIGN FTUSUF ranks above faithfulness constraints concerning stress location, except as set out in (15).

There are two other general stress constraints. The first is the three-syllable stress window at the right edge. A polysyllable may have stress on the antepenultimate syllable, for example *pel·lícula* [pəˈlikulə] 'film', or on the penultimate, for example *manicomi* [məniˈkɔmi] 'mental hospital', or on the last, for example *Algemesí* [əlʒəməˈzi] (place name). But preproparoxytone stress is absolutely excluded. The second prosodic constraint is that a word with a heavy penultimate syllable cannot have antepenultimate stress. The weight of a final syllable imposes no stress constraints, nor does the weight of an antepenultimate syllable. In (16) examples are given of all the weight possibilities in tri-syllabic words. As far as possible monomorphemic words are chosen.

The gaps in (16) consist of all and only those forms with both antepenultimate stress and heavy penults. To account for the three-syllable window together with

(16) (a) Available stress patterns in trisyllabic words with heavy penults

LHˈL	*molesquí* 'moleskin', *Novetlè* (place name)
LˈHL	*sinopsi* 'synopsis', *enorme* 'enormous'
ˈLHL	*
LHˈH	*crisantem* 'chrysanthemum', *faristol* 'lectern'
LˈHH	*adúlter* 'adulterous', *apòstol* 'apostle'
ˈLHH	*
HHˈL	*organdí* 'organdie', *almesquí* 'jonquil'
HˈHL	*extracte* 'extract', *escombra* 'broom'
ˈHHL	*
HHˈH	*restaurant* 'restaurant', *exhaustiu* 'exhaustive'
HˈHH	*escàndol* 'scandal', *extrínsec* 'extrinsic'
ˈHHH	*[12]

[12] Bonet & Lloret (1998: 205) mention ˈ*Washington* and ˈ*Amsterdam* as exceptions (lexical borrowings). But foreign place names wear their hearts on their sleeves, and are most likely to resist nativization. More indicative of active constraints is the borrowing *penalty*, pronounced in Catalan with paroxytone stress [pəˈnalti], diverging from English proparoxytone stress in exactly the way predicted by the account given here.

(b) Available stress patterns in trisyllabic words with light penults

LL'L	*comité* 'committee', *peroné* 'fibula'
L'LL	*bigoti* 'moustache', *figura* 'figure'
'LLL	*pàgina* 'page', *cèlebre* 'famous'
LL'H	*matalàs* 'mattress', *calamars* 'calamari'
L'LH	*capítol* 'chapter', *polígon* 'polygon'
'LLH	*Júpiter* 'Jupiter', *hàbitat* 'habitat'
HL'L	*Castellar* [kəstə'ʎa] (place name), *bisturí* 'scalpel'
H'LL	*misteri* 'mystery', *llenguado* 'sole'
'HLL	*àlgebra* 'algebra', *ímpetu* 'impetus'
HL'H	*enrenou* 'uproar', *actual* 'present'
H'LH	*oxigen* 'oxygen', *ventríloc* 'ventriloquist'
'HLH	*pàncrees* 'pancreas', *làudanum* 'laudanum'

the gaps in (16) I need to mention some prosodic constraints proposed in recent work. A familiar constraint is FootBinarity: Feet are binary under moraic or syllabic analysis (Kager 1999: 156). This constraint has two results: as regards the maximum size of feet, ternary and unbounded feet are disfavoured. And the minimum size of a foot is set at two moras. Hyde (2001: 58) gives good reasons for reassessing this constraint. The crucial part of it he takes to be the limit on the maximum size of a foot. For him, the bi-moraic minimality aspect falls out from other constraints. With respect to the maximum foot size, however, he proposes the FootCap Condition:

(17) FootCap Condition: Feet are maximally disyllabic.

The FootCap Condition is not a constraint (violable) but a condition on Gen; to put it another way, the FootCap Condition forms part of the definition of 'foot'. I adopt Hyde's approach here. (The minimal aspect of FtBin is in fact frequently violated in Catalan, for example in monosyllabic light prosodic words like *fe* 'faith', *pla* 'flat'. Such a foot minimality constraint is not active in Catalan grammar.)

A well-known 'parametric' constraint is the constraint selecting either trochees or iambs as the basic foot pattern. The active constraint in Catalan, as we shall see, is Iamb. The formulation in (18) follows the model for trochee given by McCarthy (2002: 150). The formulation spells out that Iamb is an alignment constraint.

(18) Iamb: AlignR(Ft, Hd(Ft)). Every foot has its head on the right-hand edge.

Iamb determines the following harmonic preferences: (L'L) ≻ ('LL), (L'H) ≻ ('LH), (H'L) ≻ ('HL), (H'H) ≻ ('HH). Monosyllabic feet, namely ('H) or ('L), do not violate Iamb as their heads are indeed aligned with the right edge of the foot (as well as the left edge).

Another familiar prosodic constraint is the Weight-to-Stress Principle (WSP; Kager 1999: 155), expressing the fact that stress preferentially falls on heavy syllables. Again I adopt here (19) Hyde's formulation (2001: 74).

(19) WHC (Weight-to-Head Constraint): Every heavy syllable must be designated as the head of some foot.

His Weight-to-Head Constraint (WHC) is in fact more general than the Weight-to-Stress Principle, as there can be heads of feet which are not stressed. Every stressed syllable must be the head of a foot, but not every head of a foot needs to be stressed.

The last crucial constraint involves the prosodic category 'colon' (κ), as argued for on the basis of data from several languages by Green (1997: 98–133), following Hayes (1995: 119) and others. The colon is a category that comes in the prosodic hierarchy below the prosodic word and above the foot, and groups feet in a manner analogous to the manner in which feet group syllables. Green (1997: 108) spells out the size of a colon with a ColonBinarity constraint (analogous to FtBin):

(20) ColonBinarity: Cola are binary (over feet and syllables).

ColonBinarity means that a colon may group two feet, such as {(σσ)(σσ)}, or a foot and a stray syllable, for example {(σσ)σ} or {σ(σσ)}. Given Hyde's argument, mentioned above, that a FootCap Condition on Gen should replace the FtBin condition, as far as the maximum size of feet is concerned, it is likely that an analogous ColonCap Condition might better account for the maximal size of cola. In fact Green considers the possibility of unbounded cola (pp. 116–19) and agrees that they are undesirable, but a violable constraint like ColonBinarity cannot rule them out in principle. I do not pursue this issue here, since the facts of Catalan do not bear directly on it. Other colon constraints that I adopt from Green (1997: 108) are spelt out in (21).

(21) (a) ParseFt: Feet are parsed into cola.
(b) AllκRt: Align(κ, R, PrWd, R). Cola are aligned with the right edge of some prosodic word. (AllκRt penalizes a word with more than one colon.)
(c) AlignHdκRt: Every head colon is aligned right with some prosodic word; i.e. 'prosodic words are right-headed' (PrWdRight in Green 1997).

In (21c) Green's PrWdRight is renamed AlignHdκRt (every head colon is aligned right with some prosodic word) as a reminder that the right-headedness of words is being evaluated by this constraint in terms not of feet but of cola.

I now turn to demonstrating how the combination (unranked with respect to one another) of AlignHdκRt (21c), Iamb (18), and WHC (19), ranked above faithfulness to input stress (FPH at (4c)), rules out first preproparoxytones and then proparoxytones with heavy penults. In (22, 23) are evaluated various

parsings and stressings of hypothetical input preproparoxytones, with H and L standing for any heavy or light syllable respectively. I ignore here the possible influence of morphological alignment constraints such as *AlignFtUSuf (12), and do not mention Head(PrWd) (4a), and ParseFt (21a), which are assumed to be undominated. Heads of feet that are not stressed are indicated in **bold**.

(22)

	'LLLL	AlignHdκRt	Iamb	WHC	FPH
(a)	{('LL)(**L**L)}			**	
(b)	{('LL)(L**L**)}			*	
(c)	{('L)}{(L**L**)L}	*			
(d)	☞ {(L'L)(L**L**)}				*
(e)	{(L)}{('LL)L)}			*	*
(f)	☞ {(L)}{(L'L)L}				*
(g)	{(L**L**)('LL)}			*	*
(h)	☞ {(L**L**)(L'L)}				*
(i)	{(L)}{(**L**L)('L)}			*	*

Candidates (22a) and (22b) show two preproparoxytone parsings that are ruled out by Iamb. Candidate (22c) avoids the Iamb violation, but at the cost of forcing the stressed syllable into a non-final colon, violating AlignHdκRt. Candidates (22d), (22f), and (22h) violate faithfulness only, but offer otherwise acceptable proparoxytone, paroxytone, and oxytone alternatives respectively, of the types *pel·lícula* 'film', *manicomi* 'mental hospital', and *Benimodí* (place name). Candidates (22e), (22g), and (22i) are less successful (trochaic) parses of the same stress patterns as (22d), (22f), and (22h).

All the preproparoxytone candidates (23a–c) violate either AlignHdκRt or Iamb. They lose to (23f), (23i), or (23l), which are the best-parsed proparoxytone, paroxytone, and oxytone candidates respectively. These correspond to the attested types *aurèola* 'halo', *pastanaga* 'carrot', and *Algemesí* (place name). Though I have only considered in (22–23) hypothetical inputs where the last three syllables are light, it should be clear that additional Hs later on will do nothing to improve the situation for preproparoxytone inputs, all of which are harmonically bound by AlignHdκRt, Iamb and WHC.

(23)

	'HLLL	ALIGNHDκRT	IAMB	WHC	FPH
(a)	{('HL)(**L**L)}		**		
(b)	{('HL)(L**L**)}		*		
(c)	{('H)}{(L**L**)L}	*			
(d)	{(H'L)(L**L**)}			*	*
(e)	{(H)}{('LL)L)}		*		*
(f)	☞ {(H)}{('L)(L**L**)}				*
(g)	{(H**L**)('LL)}		*	*	*
(h)	{(**H**L)('LL)}		**		*
(i)	☞ {(H)}{(L'L)L}				*
(j)	{(H**L**)(L'L)}			*	*
(k)	{(H)}{(**L**L)('L)}		*		*
(l)	☞ {(H)}{(**L**)(L'L)}				*

I now turn to consider trisyllabic inputs with heavy penults, in order to explain the principled gaps—those with heavy penults (*ˈσHσ)—in the (16a) set. Here I take real examples, and show in the tableaux that proparoxytone outputs are ruled out also by the constraints we have been considering, rather than by faithfulness constraints. Tableau (24) evaluates candidates for *molesquí* 'moleskin' considering various parses of the unfaithful proparoxytone, and the faithful oxytone, candidates.

In (24) each of the three proparoxytone candidates (24a–c) falls to one of the dominant constraints being considered. Candidate (24d) provides a bad parse of the correctly stressed form. Candidates (24e) and (24f) survive this evaluation. (24e) wins in the evaluation by ALLκRT, mentioned at (21b), since (24f) contains a colon which is not right-aligned in the word. Tableau (25) evaluates candidates for *sinopsi* 'synopsis'; the first three (25a–c) are identical to the failing proparoxytone candidates in (24).

(24)

	molesquí LH'L	ALIGNHDκRT	IAMB	WHC	FPH	ALLκRT
(a)	{('L)(**H**L)}		*!		*	
(b)	{('L)(H**L**)}			*!	*	
(c)	{('L)}{(H)L}	*!			*	
(d)	{(L)(H'L)}			*!		
(e)	☞ {(L**H**)('L)}					
(f)	{(L)}{(H)('L)}					*!

(25)

	sinopsi L'HL	ALIGNHDκRT	IAMB	WHC	FPH
(a)	{('L)(**H**L)}		*!		*
(b)	{('L)(H**L**)}			*!	*
(c)	{('L)}{(H)L}	*!			*
(d)	{(L)('HL)}		*!		
(e)	☞ {(L'H)L}				
(f)	☞ {(L'H)(L)}				
(g)	☞ {(L)}{('H)L}				
(h)	☞ {(L)}{('H)(L)}				

In (25e–h) four successful parses of L'HL are considered; of which (25g, h) again equally violate ALLκRT. In the following tableaux candidates with gratuitous violations of ALLκRT will not be considered. The difference between candidates (25e) and (25f) is that the final light syllable of (25e) {(L'H)L} is not parsed into a foot, while that of (25f) {(L'H)(L)}contains a violation of FTBIN (or whatever constraints produce the same effect). In the absence of evidence to choose between these I arbitrarily pick the unfooted representation of a final unstressed light syllable in the remaining tableaux, remembering that the alternative can always be substituted for it. (The constraint ALLFEETLEFT (Kager 1999: 163) also favours an unparsed word-final syllable over a degenerate foot in the same position.)

In tableau (26) I consider the two LHH patterns together (LH'H *faristol* 'lectern' and L'HH *adúlter* 'adulterous'), since the unfaithful proparoxytone 'LHH candidates are common to both.

(26)

	faristol adúlter LH'H L'HH	ALIGNHDκRT	IAMB	WHC	FPH
(a)	{('L)(**H**H)}		*!	*	*
(b)	{('L)(H**H**)}			*!	*
(c)	{('L)}{(H)(H)}	*!			*
	faristol				
(d)	{(L)(H'H)}			*!	
(e)	☞{(L**H**)('H)}				
	adúlter				
(f)	{(L)(HH)}		*!	*	
(g)	☞{(L'H)(H)}				

Candidates (26a, b, d, f) show that with IAMB and WHC constraints active there is no effective way to combine a sequence of two heavy syllables into one foot. Footing them separately, as in (26c), forces a preceding stressed syllable into a non-final colon, fatally, by ALIGNHDKRT. In (27) I examine the HHL patterns corresponding to *organdí* 'organdie' (HH'L) and *extracte* 'extract' (H'HL), repeating the model of comparison used in (26).

(27)

	organdí extracte HH'L H'HL	ALIGNHDKRT	IAMB	WHC	FPH	ALLRT
(a)	{('H)(**H**L)}		*!		*	
(b)	{('H)(H**L**)}			*!	*	
(c)	{('H)}{(H)L}	*!			*	*
	organdí					
(d)	{(H)(H'L)}			*!		
(e)	{(H**H**)('L)}			*!		
(f)	{(**H**H)('L)}		*!	*		
(g)	☞ {(H)(H)}{('L)}					*
(h)	☞ {(H)}{(H)('L)}					*
	extracte					
(i)	{(H)('HL)}		*!			
(j)	☞ {(H)}{('H)L}					*

In (27a, b) a final HL sequence cannot, subject to IAMB and WHC, compose a single foot; hence, as in (27c), a preceding stressed foot has to be in a separate, non-final colon, violating ALIGNHDKRT. The comparison between candidates (27c) and (27j), by the way, shows why I reject the view that the trochee is the unmarked foot pattern in Catalan. (HL) would form a good trochee; therefore {('H) (HL)} would form a good bi-pedal head colon. In (27d–f) various disyllabic foot parses with an initial heavy syllable violate IAMB or WHC or both. Candidates (27g–h), equally faithful to input stress, are evaluated by ALLKRT which, if evaluation is categorical, does not choose between them—there is no difference in pronunciation. In (27i) it is shown again that a disyllabic foot with a heavy initial does not pass IAMB.

Lastly, in (28) I consider HHH candidates of the types *restaurant* 'restaurant' and *extrínsec* 'extrinsic'.

(28)

	restaurant extrínsec HH'H H'HH	ALIGNHDKRT	IAMB	WHC	FPH	ALLKRT
(a)	{('H)(**H**H)}		*!	*	*	
(b)	{('H)(H**H**)}			*!	*	
(c)	{('H)}{(H)(H)}	*!			*	*
	restaurant					
(d)	{(H)(H'H)}			*!		
(e)	{(H**H**)('H)}			*!		
(f)	{(**H**H)('H)}		*!	*		
(g)	☞ {(H)(H)}{('H)}					*
(h)	☞ {(H)}{(H)('H)}					*
	extrínsec					
(i)	{(H)('HH)}		*!	*		
(j)	☞ {(H)}{('H)(H)}					*

The proparoxytone candidates (28a–c) are ruled out, as previously, by high-ranking ALIGNHDKRT, IAMB, and WHC; and candidates (28g)/(28h) for *restaurant*, and (28j) for *extrínsec*, win over the considered alternatives.

All remaining patterns of prosodic words are well formed in Catalan. Table (16b), repeated here as (29), illustrated the range of trisyllables with light penults. Table (30) illustrates the range of disyllables, and (31) the monosyllabic prosodic words. The foot and colon parses in the right-hand columns are those that pass the constraints ALIGNHDKRT, IAMB, and WHC which have been illustrated previously. In the appendix to this chapter, these successful parses are evaluated against a number of alternatives.

(29) Available stress patterns in trisyllables with light penults

LL'L	*comité* 'committee', *peroné* 'fibula'	{(L)(L'L)} or {(L**L**)('L)}*
L'LL	*bigoti* 'moustache', *figura* 'figure'	{(L'L)L}
'LLL	*pàgina* 'page', *cèlebre* 'famous'	{('L)(L**L**)}
LL'H	*matalàs* 'mattress', *calamars* 'calamari'	{(L)(L'H)} or {(L**L**)('H)}*
L'LH	*capítol* 'chapter', *polígon* 'polygon'	{(L'L)(H)}
'LLH	*Júpiter* 'Jupiter', *hàbitat* 'habitat'	{('L)(L**H**)}
HL'L	*Castellar* [kəstə'ʎa] (place name), *bisturí* 'scalpel'	{(H)(L'L)}
H'LL	*misteri* 'mystery', *llenguado* 'sole'	{(H)}{('L)L}
'HLL	*àlgebra* 'algebra', *ímpetu* 'impetus'	{('H)(L**L**)}
HL'H	*enrenou* 'uproar', *actual* 'present'	{(H)(L'H)}
H'LH	*oxigen* 'oxygen', *ventríloc* 'ventriloquist'	{(H)}{('L)(H)}
'HLH	*pàncrees* 'pancreas', *làudanum* 'laudanum'	{('H)(L**H**)}

*The first analysis is favoured by ALIGN(L)Ft = ALLFEETLEFT.

(30) Available stress patterns in disyllabic words

L'L	*cafè* 'coffee', *menú* 'table d'hôte'	{(L'L)}
'LL	*tribu* 'tribe', *mare* 'mother'	{('L)L}
L'H	*segon* 'second', *sotrac* 'jolt'	{(L'H)}
'LH	*llapis* 'pencil', *pèsol* 'pea'	{('L)(H)}
H'L	*garbí* 'south west wind', *tauró* 'shark'	{(H)('L)}
'HL	*ardu* 'arduous', *taxi* 'taxi'	{('H)L}
H'H	*engany* 'deceit', *pinzell* 'paintbrush'	{(H)('H)}
'HH	*càstig* 'punishment', *príncep* 'prince'	{('H)(H)}

(31) Available stress patterns in monosyllabic words

'H	*tros* 'piece', *verd* 'green'	{('H)}
'L	*cru* 'raw', *re* 'anything'	{('L)}

9.6 RHYTHMIC CONSTRAINTS ABOVE THE WORD: A PROVISIONAL ACCOUNT

Let us start with some simplifying assumptions. I assume here that the relevant prosodic category above the phonological word is the phonological phrase (PPhr). I make only a few brief observations concerning the integration of phonological phrases into intonation phrases. Secondly, I assume that phonological phrases are built out of prosodic words (or lexical words—see below), and that the relevant correspondence relation is between the form of independent words and their correspondents in PPhrs. The logic of this position would lead one to use the same terminology, i.e. 'stressed' and 'unstressed', at both levels. At the word level, however, the notions of 'stressed' and 'unstressed' are strongly associated with other lexical phenomena, especially the demarcative function of stress—one only per phonological word—and the associated phenomenon of vowel reduction (§2.3; see also Beckman 1998; Alderete 1999: 36). At the PPhr level, prominence and its absence, though driven primarily by PrWd stress, are associated with different phenomena. Thus, for example, the phrase *tots tres junts* 'all three together', which consists of three prosodic words which are stressed monosyllables, is likely to be realized with greater prominence on the first (*tots*) and last (*junts*), with the intervening *tres* being rhythmically weak. Here I use 'strong' (S) and 'weak' (W) for the high and low points (syllables) of rhythmic prominence. Thus I represent the rhythm of *tots tres junts* as SWS, in which stressed *tres* is rhythmically W(eak). In the phrase *la* ˈ*màxima felici*ˈ*tat* 'the greatest happiness' there are two PrWd stresses, but a likely pronunciation would show three peaks of rhythmic prominence (indicated in bold): *la* ˈ***mà**xima* ***fe**lici*ˈ***tat***, corresponding to a rhythmic template WSWWSWWS. Here the first syllable of *felicitat* is unstressed, but rhythmically S(trong). The correspondence

constraint (which is violated once in each of the two phrases considered so far) is formulated in (32).

(32) IDENT$_{\text{LexWd-PPhr}}$Prosodic prominence (IDENT$_{\text{LP}}$Prom): Prosodic prominence in the lexical word is identical to prosodic prominence in the phonological phrase. (LexWd 'stressed' is identical to PPhr 'Strong' and LexWd 'unstressed' is identical to PPhr 'Weak'.)

In (32) IDENT$_{\text{LP}}$Prom is formulated in terms of the lexical word, since we need to include clitics, which are lexical but not prosodic words (PrWds) and are thus unstressed. The first example above, *tots tres junts* SWS, violates IDENT$_{\text{LP}}$Prom in order to satisfy the rhythmic constraint *CLASH, while the second, *la màxima felicitat* WSWWSWWS, has one violation of IDENT$_{\text{LP}}$Prom in order to satisfy the rhythmic constraint *LAPSE. *CLASH and *LAPSE are the major constraints driving rhythmic alternation. *CLASH penalizes SS, while *LAPSE penalizes WWW (33). (Other members of the *CLASH and *LAPSE families are illustrated in §3.2 and §4.1.)

(33) (a) *CLASH: *SS. No adjacent Strong syllables.
(b) *LAPSE: *WWW. No more than two adjacent Weak syllables.

There are constraints (34) governing the headedness and the alignment of phonological phrases which are analogous to the ones at (4) governing prosodic words.

(34) (a) HEAD(PPhr): Each phonological phrase has a unique head.
(b) ENDRULE-R(PrWd): The head prosodic word is not followed by another prosodic word in the phonological phrase.

The constraints HEAD(PPhr) and ENDRULE-R(PrWd) are undominated in the relevant rankings. In what follows, I assume this, and do not consider candidates that violate IDENT$_{\text{LP}}$Prom at the right edge (as would *tots tres junts* *WSW, for example).

As should be clear from the discussion and examples mentioned so far, *CLASH and *LAPSE dominate IDENT$_{\text{LP}}$Prom. I illustrate this in tableau (35) with the phrases *aquells tres amics* 'those three friends' and *aquells tres carnissers* 'those three butchers' (Oliva 1992: 67-9).

(35)

	a'quells 'tres a'mics	*CLASH	*LAPSE	IDENT$_{\text{LP}}$Prom
(a)	W S S W S	*!		
(b)	☞W W S W S			*
(c)	☞W S W W S			*
(d)	S W W W S		*!	***

	a'quells 'tres carni'ssers	*CLASH	*LAPSE	IDENT$_{LP}$Prom
(e)	W S S W W S	*!		
(f)	W S W W W S		*!	*
(g) ☞	W W S W W S			*
(h)	W S W S W S			**

Candidates (35b) and (35c) both avoid *CLASH and *LAPSE, and are both well-formed. In (35f), however, the same rhythm as in (35c) for *aquells tres* leads to a *LAPSE violation, unless, as in (35h) the first syllable of *carnisser* becomes S. Though both (35g) and (35h) pass *CLASH and *LAPSE, (35h) has more violations of IDENT$_{LP}$Prom, and on this basis ought to lose. We shall see below that the issue is a little more complicated.

Consider now the phrase *di'sset 'nenes finlan'deses* 'seventeen Finnish girls', which, on the basis of PrWd stresses, displays one clash and one lapse. Several conceivable alternatives (not ruled out by *CLASH and *LAPSE) are listed in tableau (36).

(36)

	di'sset 'nenes finlan'deses	*CLASH	*LAPSE	IDENT$_{LP}$Prom	ALIGNSRT
(a)	W S S W W W S W	*!	*		
(b)	W S W W S W S W			**	1 3 6 (10)!
(c) ☞	W W S W S W S W			**	1 3 5 (9)
(d) 💣	S W S W S W S W			***!	1 3 5 7 (16)
(e)	W S W S W W S W			**	1 4 6 (11)!

Candidates (36b–e) all avoid *CLASH and *LAPSE, but, in fact, (36c) and (36d) alone are acceptable (see below for remarks on (36d)). A relevant difference between (36b) and (36c) (though not the only one, as will be seen) is that (36c) has its first S one syllable further right than (36b); their other Ss match. So (36c) wins on ALIGNSRT, where violations are counted according to how many syllables intervene between each S and the right edge (totals in parentheses). Candidate (36e) also loses to (36c) by this criterion, but also displays a problem with how *'nenes* is realized: WS. Here, the prominence reversal looks problematic, though, as we shall shortly see with (36d), reversal itself is not lethal. What is wrong with *'nenes* realized WS is that a PrWd post-tonic syllable is realized as S in the phonological phrase. It is plausible that the demarcative function of word stress is compromised when prominence is realized further right than it 'should'

be. A correspondence constraint more specific than IDENT_{LP}Prom is involved, which I here (37) abbreviate as *INVERTPromRt.

(37) *INVERTPromRt: The head of a lexical word and a following unstressed syllable are not inverted (to Weak + Strong) in the phonological phrase.
Note that only SW → WS is penalized; SW may be realized WW, WS may be realized SW or WW, and SWW may be realized SWS without penalty. The problem is prominence reversal with prominence moving rightwards.

I have formulated the constraint *INVERTPromRt in terms of the lexical word bearing in mind Oliva's observation (1992: 89) of the rhythmic difference between ˈ*tres* ˈ*portes se*ˈ*cretes* 'three secret doors' (WSWWSW) and ˈ*tres* ˈ*portamo*ˈ*nedes* 'three purses' (SWSWSW). In the latter, which contains a lexical compound of the prosodic words ˈ*porta* 'carries' and *mo*ˈ*nedes* 'coins' where the head syllable of the lexical word is [ˈnɛ], the stressed/unstressed pattern of the first component ˈ*porta* may be realized WS, a rhythmic pattern which is not available to *portes* 'doors', by *INVERTPromRt. The ranking of *INVERTPromRt is a matter to be considered shortly. By the constraints considered so far, candidate (36d), realizing *di*ˈ*sset* ˈ*nenes finlan*ˈ*deses* on the (binary trochaic) pattern SWSWSWSW, ought to be ungrammatical (hence the 💣 in (36)). In fact (Oliva 1992: 65–6) it is an acceptable alternative to (36c) WWSWSWSW, despite the fact that it bears an extra S (maximally distant from the right edge), and despite the fact that the iambic pattern of lexical *di*ˈ*sset* is realized with reversed prominence SW. What the acceptable candidate (36d) illustrates is an optional preference for the PPhr to begin with a prominent (S) syllable, a constraint formulated in (38).

(38) INITProm: Every phonological phrase begins with a prominent syllable (S).

The comparison of (36c) and (36d) shows that, when active, INITProm is ranked above IDENT_{LP}Prom. Phrase-initial SS, though, is just as intolerable a clash as elsewhere, so INITProm ranks below *CLASH. Examination (39) of the phrase *les plàcides paraules* 'the calm words' (WSWSWSW) will show that the correspondence constraint IDENT_{LP}Prom needs to be divided into MAX and DEP components, where MAX penalizes realizing a word stress as rhythmically Weak, while DEP penalizes realizing a lexically unstressed syllable as rhythmically Strong. INITProm ranks between MAX_{LP}Prom and DEP_{LP}Prom, as will be explained.

(39)

	les 'plàcides pa'raules	*CLASH	*LAPSE	MAX$_{LP}$Prom	(INITProm)	DEP$_{LP}$Prom	ALIGNSRT
(a)	W S W W W S W		*!		*		1 5 (6)
(b)	☞ W S W S W S W				*	*	1 3 5 (9)
(c)	W W S W W S W			*!	*	*	1 4 (5)
(d)	S W S W W S W			*!		**	1 4 6 (11)
(e)	S W W S W S W			*!		**	1 3 6 (10)

In tableau (39) candidate (39b) wins, whether or not INITProm is active. If INITProm is not active, (39b) has no MAX violations (better than (39c–e)). ALIGNSRT is irrelevant. If INITProm is active, however, then, if MAX$_{LP}$Prom ranked lower than it, only candidates (39d) and (39e) would survive, to be evaluated by ALIGNSRT, whereby (39e) would wrongly be the winner. With MAX$_{LP}$Prom » INITProm, as required, candidates that lose on MAX$_{LP}$Prom are not in the running.

The true lesson from the comparison of (36c) and (36d) is that INITProm ranks above DEP$_{LP}$Prom. This is illustrated in tableau (40) with the phrase *pastanagues bullides* 'boiled carrots'. Two versions are given, first with INITProm active (40a: SWSWWSW) and then without INITProm (40b: WWSWWSW). In each, candidates (iii, iv) with gratuitous 'destressing' of *pastanagues* are knocked out by high-ranking MAX$_{LP}$Prom.

(40) (a)

	pasta'nagues bu'llides	*CLASH	*LAPSE	MAX$_{LP}$Prom	INITProm	DEP$_{LP}$Prom	ALIGNSRT
i.	W W S W W S W				*!		1 4
ii.	☞ S W S W W S W					*	1 4 6
iii.	S W W S W S W			*!		**	1 3 6
iv.	W S W S W S W			*!	*	**	1 3 5

(b)

pasta'nagues bu'llides	*Clash	*Lapse	MaxLPProm	DepLPProm	AlignSRt
i. ☞ W W S W W S W					1 4
ii. S W S W W S W				*!	1 4 6
iii. S WW S W S W			*!	**	1 3 6
iv. W S W S W S W			*!	**	1 3 5

In version (40a), candidate (40a.i), with faithful initial WW, falls to InitProm, and the remaining candidate, (40a.ii), with the 'repaired' initial SW, is the victor. In version (40b), without the constraint InitProm, it is the faithful candidate (40b.i) that wins. In (39) and (40) I had no occasion to call on the constraint *InvertPromRt, even though (39b–e) and (40a/b.iii, iv) illustrate violations of it, since these candidates were ruled out anyway by other constraints. The phrase *'deu 'plàcides pa'raules* 'ten calm words' (WSWSWSW), which is similar to *les 'plàcides pa'raules* of (39), illustrates that *InvertPromRt (*Inv) must rank above InitProm, as is displayed in tableau (41a). Tableau (41b) shows the alternative with InitProm inactive.

(41) (a)

'deu 'plàcides pa'raules	*Clash	*Lapse	MaxLPProm	*Inv	InitProm	DepLPProm	AlignSRt
i. S S W W W S W	*!	*					1 5 6 (12)
ii. ☞ W S W S W S W			*		*	*	1 3 5 (9)
iii. W W S W W S W			**!	*	*	*	1 4 (5)
iv. S W S W W S W			*	*!		*	1 4 6 (11)
v. S W W S W S W			*	*!		*	1 3 6 (10)

(b)

'deu 'plàcides pa'raules	*Clash	*Lapse	MaxLPProm	*Inv	DepLPProm	AlignSRt
i. S S WW W S W	*!	*				1 5 6 (12)
ii. ☞ W S WS W S W			*		*	1 3 5 (9)
iii. W W SW W S W			**!	*	*	1 4 (5)
iv. S W SW W S W			*	*!	*	1 4 6 (11)
v. S WW S W S W			*	*!	*	1 3 6 (10)

If InitProm were ranked higher than *InvertPromRt in (41a), candidate (41ii) would be eliminated. Evaluation would then pass to *InvertPromRt, on which (41iv) and (41v) tie, as they do on Dep_{LP}Pr. Evaluation would pass to AlignSRt by which (41v) would win, which is incorrect. In fact, (41ii) WSWSWSW is the only acceptable version, whether InitProm is active or not.

I remarked above that *InvertPromRt penalizes realizing SW within a word as WS, but not other failures of correspondence. Some examples follow illustrating this claim. First consider (42) the phrase *amb a'quell 'bon com'pany* 'with that good friend' (Oliva 1992: 83). As before, I give two versions, with InitProm (42a: SWWSWS) and without (42b: WWSWWS).

(42)

(a)

amb a'quell 'bon com'pany	*Clash	*Lapse	Max_{LP}Prom	*Inv	InitProm	Dep_{LP}Prom	AlignSRt
i. W W S S W S	*!				*		2 3 (5)
ii. W W S W W S			*		*!		3
iii. W S W S W S			*		*!	*	2 4 (6)
iv. ☞ S W W S W S			*			*	2 5 (7)
v. S W S W W S			*			*	3 5 (8)!

(b)

amb a'quell 'bon com'pany	*Clash	*Lapse	Max_{LP}Prom	*Inv	Dep_{LP}Prom	AlignSRt
i. W W S S W S	*!					2 3 (5)
ii. ☞ W W S W W S			*			3
iii. W S W S W S			*		*!	2 4 (6)
iv. S W W S W S			*		*!	2 5 (7)
v. S W S W W S			*		*!	3 5 (8)

In (42a), candidates (42a.ii, iii) are discarded by InitProm. Candidates (42a.iv) and (42a.v) tie on both Max_{LP}Prom and Dep_{LP}Prom, and evaluation passes to AlignSRt, by which (42a.iv) is the winner. In (42b), without InitProm, candidates (42b.iii–v) fall to Dep_{LP}Prom, and (42b.ii) is the winner. Observe that in (42b.iv) the lexical iamb *a'quell* is acceptably realized WW.

In (43) consider the phrase *el ca'rrer 'Gran* 'the High Street', again in two versions, with (SWWS) and without active InitProm (WSWS).

(43) (a)

el ca'rrer 'Gran	*Clash	*Lapse	MaxLPProm	*Inv	InitProm	DepLPProm	AlignSRt
i. W W S S	*!				*		1
ii. W S W S			*		*!	*	2
iii. ☞ S W W S			*			*	3

In (43a) with active InitProm, it proves acceptable to realize *ca'rrer* as WW. In (43b) candidate (43b.ii) wins on AlignSRt; note how *ca'rrer* is realized as SW ('inverted', with prominence moving left).

(b)

el ca'rrer 'Gran	*Clash	*Lapse	MaxLPProm	*Inv	DepLPProm	AlignSRt
i. W W S S	*!					1
ii. ☞ W S W S			*		*	2
iii. S W W S			*		*	3!

Lastly I review a couple of phrases where *Clash is not involved, and therefore MaxLPProm is not in play either (as only a clash could be improved by a Max violation realizing a word stress as W). The phrases are *'moltes felici'tats* 'many congratulations' (SWWSWS) (44) and *'màximes infelici'tats* 'greatest misfortunes' (SWWSWSWS) (45).

In (44) and (45) the issue is decided by the AlignSRt constraint, which favours binary stress alternations starting from the right, other things being equal.[13]

(44)

'moltes felici'tats	*Clash	*Lapse	MaxLPProm	*Inv	InitProm	DepLPProm	AlignSRt
S WWWW S		**!					5
S W SWW S						*	3 5 (8)!
☞ S W WSW S						*	2 5 (7)

[13] I have at present no secure explanation of why, according to Oliva (1992: 78), in *'és de les que s'i'rriten* 'she is one of those who gets annoyed', SWSWWSW is more eurhythmic than SWWSWSW (which by AlignSRt is better aligned). I suspect it has to do with the specific morphemes involved:

(45)

ˈmàximes infeliciˈtats	*CLASH	*LAPSE	MAX$_{LP}$Prom	*INV	INITProm	DEP$_{LP}$Prom	ALIGNSRT
S WWWWWW S		****!					7
☞ S WW SWSW S						**	2 4 7 (13)
S W S WWSW S						**	2 5 7 (14)!
S W S WSWW S						**	3 5 7 (15)!

The matter of the syntactic scope of the phonological phrase is by no means straightforward. Roughly speaking, a phonological phrase may consist of a noun phrase and its complements, a prepositional phrase or an adverbial, or a verb group (auxiliary, lexical verb and associated adverbs, such as *ˈno havˈia ˈvistˈ mai* 'had never seen' SWSWWS, or, with elision of the first syllable of *havia*, WSWWS). On this basis the sentence (one intonation phrase) *en ˈPep ˈdorm a la ˈplatja* 'Pep is asleep on the beach' (Oliva 1992: 102) consists of three phonological phrases [*en ˈPep*]$_{NP}$ [*ˈdorm*]$_{V}$ [*a la ˈplatja*]$_{PP}$. The constraints HEAD(PPhr) and ENDRULE-R(PrWd) (34) imply (or must be interpreted to mean) that in an intonation phrase like *en ˈPep ˈdorm a la ˈplatja* none of the heads can be demoted with the goal of removing or improving an apparent clash—WSSWWSW in this example. However, as Oliva illustrates (pp. 103–4), it is permissible, though not obligatory, to combine a verb and its complement(s) (i.e. a VP) into one PPhr, as *ˈdorm ˈsense piˈjama* 'sleeps without pyjamas', realized WSWWSW. It is also permissible to combine a subject NP and a verb into one PPhr, as in *en ˈPep ˈdorm* 'Pep is asleep' WWS. The constraints on such PPhr combining, which clearly involve both syntactic aspects like the relation between adjacent phrases[14] and phonological ones like the length of the phrases, require further research. It is apparent, though, that within intonation phrases, clashes and lapses that arise from adjacent phonological phrases are not unacceptable.

les, although a clitic, is head of its DP/NP, and semantically more prominent than *que*, the relative particle or complementizer. However, Oliva claims (p. 100) that SWSWWS is *generally* more eurhythmic than SWWSWS; see (422) above.

[14] So the ambiguous string *un baˈrret de ˈnen ˈnegre* 'a [black [child's hat]]' or 'a [[black child's] hat]' may have two different rhythms depending on the syntax.

APPENDIX: EVALUATION OF ALTERNATIVE FOOT AND COLON PARSES FOR WORD TYPES IN (29–31)

comité LL'L	ALIGNHDKRT	IAMB	WHC
☞ {(L)(L'L)}			
☞ {(L**L**)('L)}			
{(**L**L)('L)}		*	
bigoti L'LL			
☞ {(L'L)L}			
{(L)('LL)}		*	
cèlebre 'LLL			
☞ {('L)(L**L**)}			
{('LL)L}		*	

matalàs LL'H	ALIGNHDKRT	IAMB	WHC
☞ {(L)(L'H)}			
☞ {(L**L**)('H)}			
{(**L**L)('H)}		*	
capítol L'LH			
☞ {(L'L)(H)}			
{(L)('LH)}		*	*
Júpiter 'LLH			
☞ {('L)(L**H**)}			
{('LL)(H)}		*	

bisturí HL'L	ALIGNHDKRT	IAMB	WHC
☞ {(H)(L'L)}			
{(H**L**)('L)}			*
{(**H**L)('L)}		*	
misteri H'LL			
{(H'L)L}			*
{(H)('LL)}v		*	
☞ {(H)}{('L)L}			
ímpetu 'HLL			
☞ {('H)(L**H**)}			
{('HL)(H)}		*	

actual HL'H	ALIGNHDKRT	IAMB	WHC
☞ {(H)(L'H)}			
{(H**L**)('H)}			*
{(**H**L)('H)}		*	
oxigen H'LH			
{(H'L)(H)}			*
{(H)('LH)}		*	*
☞ {(H)}{('L)(H)}			
pàncrees 'HLH			
{('H)}{(L)(H)}	*		
☞ {('H)(L**H**)}			
{('HL)(H)}		*	

menú L'L	AlignHdkRt	Iamb	WHC	FtBin
☞{(L'L)}				
{(L)('L)}				*
mare 'LL				
{('LL)}		*		
☞{('L)L}				*
segon L'H				
☞{(L'H)}				
{(L)('H)}				*
llapis 'LH				
{('LH)}		*	*	
☞{('L)(H)}				*

tauró H'L	AlignHdkRt	Iamb	WHC	FtBin
{(H'L)}			*	
☞{(H)('L)}				*
taxi 'HL				
{('HL)}		*		
☞{('H)L}				
engany H'H				
{(H'H)}			*	
☞{(H)('H)}				
càstig 'HH				
{('HH)}		*	*	
☞{('H)(H)}				

verd 'H	ALIGNHDκRT IAMB WHC	FTBIN
☞{('H)}		
cru 'L		
☞{('L)}		*

10

WORD PHONOLOGY AND PHONOLOGICALLY CONDITIONED ALLOMORPHY

10.1 CODA/ONSET ALTERNATIONS: LENITION

10.1.1 *Introduction*

Catalan participates in the areal phenomenon (along with Spanish, Basque, and some Occitan varieties) of voiced obstruent stop lenition. In Catalan the pattern usually described consists of voiced fricatives (or approximants) in onsets after continuant segments ([+cont]), and voiced plosives in all other contexts (see details of these in (8)). In simple terms, /b/ has allophones [b] and [β], /d/ has allophones [d̪] and [ð], and /g/ has allophones [g] and [ɣ]. This allophonic distribution, traditionally labelled 'lenition', has received some attention in recent work on Catalan phonology, without, however, being the object of a detailed account. To introduce the analysis and interpretation offered here I set out a list (1) of eight issues in Catalan lenition that have arisen in treatments of Catalan phonology.

(1) i. Are the continuant allophones fricatives or approximants? In the Catalan linguistic tradition they were regarded as fricatives till Recasens (1986: 86–90) suggested that they were more appropriately described as approximants. Since then Recasens's terminology has been followed. I believe Recasens's interpretation needs re-examination. This is the theme of §10.1.2.

ii. Is Catalan lenition a phonetic or a phonological process? The question concerns the theoretical issue of which sub-domain of theory offers the right explanation of the phenomenon. Mascaró (1991) defends the phonetic character of the process, while Palmada (1997) takes the opposing view.

iii. Although generally the fricative variants are found between continuant segments ([+cont]), there are some apparent exceptions. For example, the dental stop rather than the fricative is found after a lateral, as in *alcalde* [əlˈkaldə] 'mayor', *ull de bou* [uʎ də ˈβɔw] 'round window'. It is inadequate to adduce merely homorganicity (or approximate homorganicity) of the preceding consonant as blocking lenition, given that the dental fricative occurs after an alveolar sibilant, as in *esdevenir*

[əzðəβəˈni] 'become'. It is a matter of establishing an adequate and well-motivated definition of the leniting context. Similarly the bilabial plosive, not fricative, is found after a labiodental fricative as in *bolígraf boníssim* [buˈliɣɾəv buˈnisim] 'excellent ballpoint'. This particular 'exception', involving a rather infrequent consonant sequence (/fb/), is satisfactorily explained by Recasens (1993: 181, 187) in terms of the articulatory complexity of the sequence. I do not consider it specifically here.

iv. The alternation between the voiced alveolo-palatal affricate and the fricative, as in *fuig* [ˈfutʃ] (with word-final devoicing) 'flee.2SG.IMP' – *fugim* [fuˈʒim] 'flee.1PL.PRS', displays some, but not exact, parallelism with the phenomenon of voiced plosive lenition. The question arises whether the parallel phenomena should be treated as aspects of the same process, as is proposed by Mascaró (1984b: 215).

v. Catalan lenition resembles the phenomenon reported in Spanish and Basque, but there are significant differences, as is indicated by the pronunciation of the word *abdicar* 'abdicate.INF': Catalan [əbdiˈka], Spanish [aβðiˈkaɾ]. Catalan lenition requires a [+cont] segment on both sides, whereas Spanish (and Basque) lenition requires only a preceding [+cont] segment. This theme of cross-linguistic typological variation is considered by Mascaró (1984b) and Hualde (1991: 99–107).

vi. There is some regional and social variation in Catalan lenition. This fact is ignored in the majority of phonological studies, though it is mentioned briefly by Mascaró (1984b; 1991: 173–4). It is covered more extensively by Recasens (1993, and particularly 1991b: 183, 221, 234). It is noted that the bilabial fricative [β] allophone of /b/ is less common in those dialects that retain the voiced labio-dental fricative phoneme /v/. Thus there are both sociolinguistic and structural factors in play.

vii. What kind of process is lenition? Within Generative Phonology lenition is assumed, with little or no discussion, to fall under the heading of assimilation, or 'spreading' of the feature [+cont] (Mascaró 1984b; 1991; Palmada 1994a; 1994b; 1995; 1997; Bonet & Lloret 1998: 156–60). But lenition can equally, or better, be considered a weakening process. Of course, weakening processes were hard to capture or express with the mechanisms of classical Generative Phonology.

viii. What should be the underlying representation of the phonemes with stop and fricative allophones? Phonologists dealing with Catalan have generally taken the position, usually with little argument, that the stop variant is underlying. Only Hualde (1991: 106–7) proposes the contrary. More exactly, he proposes that the underlying forms are unspecified for the feature [continuant], while a default rule adds [+cont] to the underspecified /B, D, G/.

The issues mentioned in (1) are not the only ones a phonologist might be concerned with. There are yet more basic ones that have been generally ignored.

Why does lenition exist? Why does it affect in particular voiced obstruents, giving non-strident stops and fricatives (or approximants) as allophones? What is the reason for its occurrence in some contexts but not others? I address these basic questions here, while also considering, in part at least, the issues already found in the literature and mentioned in (1). One of the most important consequences of Optimality Theory is that it drives researchers to address functional questions. In addition to establishing what happens, in what contexts, one is obliged to ask: why does a phonological process (such as lenition) exist? Why does it occur in certain contexts and not in others? Why does it result in a particular type of output and not in others? Functional answers are typically given in terms of physiological or psychological explanations. Physiological explanations bring in aspects of articulatory phonetics or aerodynamics, while psychological explanations may involve, for example, psycho-acoustics or matters concerning speakers' communicative intentions.

10.1.2 *Fricatives or approximants*

As mentioned in (1.i), the non-stop allophones of voiced non-strident obstruent phonemes were described as fricatives until 1986, when Recasens (1986: 86) argued for their more exact characterization as approximants. He maintained that acoustic investigation demonstrated that these consonants did not display the turbulence characteristic of typical fricatives like sibilants [s], [ʃ], or the labio-dental [f]. But he did not in fact demonstrate that the non-strident continuant obstruents positively displayed properties characteristic of the class of approximants. And not all phoneticians make use of the term 'approximant', which denotes, in essence, a phonological rather than a phonetic category. Those who do use 'approximant' (Ladefoged 2000) use it for consonants (sonorants, in fact) that possess many vowel-like qualities, such as semivowels, and among many phonologists and phoneticians, all rhotics and (voiced) laterals.[1] But in a somewhat different phonetic approach, which is represented, for example, in Kirchner (1998), fricatives may be of various kinds. In some types of fricative there is turbulence and high-frequency noise; these are strident fricatives, the typologically unmarked fricatives, including sibilants and labio-dentals. But other fricatives are non-strident, less noisy, typologically less frequent and more marked, including bilabials, non-sibilant dentals, and velars ('spirants' in more traditional terminology). The typological markedness relations among fricative subtypes are seen in the fact that a language may have strident fricatives without non-strident ones (such as French or Italian), but no language has non-strident fricatives

[1] Clements (1990: 293) gives a definition, 'any sound produced with an oral tract stricture open enough so that the airflow through it is turbulent only if it is voiceless', that seems to allow the inclusion of voiced non-stridents like [β], [ð], [ɣ], if non-turbulent. But he discusses only 'liquids, glides and vowels', and in the context of his discussion of sonority clearly does not intend to include any obstruents. Kenstowicz (1994: 34, 255) refers only to glides and liquids.

without some strident fricatives also. Kirchner (1998: 99) shows that the lenition of stops always results in non-strident fricatives. In feature terms, the weakening of a non-strident stop does not alter the value for [strident], but affects only continuance. As far as Catalan is concerned, the bilabial, dental, and dorsal continuants behave phonologically like obstruents, i.e. like fricatives, not like sonorants, which would be expected if they were approximants.

One context where the obstruent-like properties of non-strident continuants is observed is in the Minimum Sonority Distance constraint which regulates the syllabification of consonants in Catalan (along with the more general Sonority Sequencing principle SonSeq; §3.1). Versions of Minimum Sonority Distance (MSD) have been discussed in §3.1 (6) and §8.2 (4, 5). The sonority ranking from Chapter 8 is repeated here as (2) and the MSD6 constraint (§8.2 (5)) is repeated here as (3).

(2) Sonority ranking (Catalan)

15	9	8	5	4	2	0
Glides	Tap	Laterals, trill	Nasals	Sibilant continuants	Non-strident continuant obstruents	Non-sibilant stops/strident continuants
j w	ɾ	l ʎ r	m n ɲ	s ʃ z ʒ	β ð ɣ	f p t k b d g

(3) MSD6: In an onset sequence C_1C_2, the value of $C_2 - C_1 \geq 6$

Examples:

V.frV, V.prV, V.trV, V.krV, Vm.brV, Vn.drV, Vɲ.grV	[$_\sigma$0-9V
V.flV, V.plV, V.klV, Vm.blV, Vŋ.glV	[$_\sigma$0-8V
V.βrV, V.ðrV, V.ɣrV	[$_\sigma$2-9V
V.βlV, V.ɣlV	[$_\sigma$2-8V
*V.βʒV, *V.ðzV, *V.ɣzV	*[$_\sigma$2-4V
*V.slV, *V.srV	*[$_\sigma$4-9V
*V.lrV	*[$_\sigma$8-9V
V.frjV, V.prwV, V.trjV, V.krwV, Vm.brjV, Vn.drwV, Vn.grjV	[$_\sigma$0-9-15V
V.fljV, V.pljV, V.kljV, Vm.bljV, Vŋ.gljV	[$_\sigma$0-8-15V
V.βrjV, V.ðrjV, V.ɣrjV	[$_\sigma$2-9-15V
V.βljV, V.ɣljV	[$_\sigma$2-8-15V

Observe that in the syllabification of complex onsets, non-sibilant obstruent continuants behave just like other non-sibilant obstruents, and in no way like semivowels or liquids (true 'approximants'). For example, neither sibilant fricatives nor nasals make complex onsets with liquids because they are too close on the sonority gradient. Non-sibilant fricatives, however, being even lower than sibilants on the sonority gradient, can make complex onsets with liquids. On this basis I return to the earlier terminology, and speak of fricative allophones of

voiced (non-sibilant) obstruent phonemes, while recognizing that these are weak non-strident fricatives.

10.1.3 *Lenition and the theory of effort*

In order to understand the motivation for the process of consonant lenition, investigations of articulatory effort, such as those of Kirchner (1998) or Boersma (1998), are particularly helpful. I discuss here only Kirchner's account, which is directed specifically at the lenition phenomenon. Kirchner argues that lenition reveals the activity of a constraint that can be summarized as in (4).

(4) LAZY: Minimize articulatory effort (Kirchner 1998: 38).

Lenition is the outcome when the LAZY constraint dominates some faithfulness constraint, such as IDENT(cont), which would require the output value of [continuant] to match the input value. 'Effort' refers to neuromuscular or biomechanical activity, and is a function of mass, velocity, and time. In principle effort can be quantified, and both Kirchner and Boersma develop mathematical models of effort in consonant articulation. For the analysis of Catalan consonant variants given here, the important thing is to develop the concept of *relative* effort, whereby one articulation can be shown to be more effortful than another comparable one.

Pursuing the idea that effort is a function of mass, velocity, and time, it follows that larger or more rapid gestures require more effort. Articulatory precision would also be an activity reducible to notions of effort, in this case of a neuromuscular rather than a biomechanical kind. Thus, for example, the articulation of a sibilant fricative requires precision of effort in order to resist the upward movement of the articulator, and to realize a narrow constriction without blocking the airflow completely. Some articulations require more precision than others. Generally speaking, a plosive is articulated with a simple ballistic gesture, for which it is sufficient that the effort of displacement towards the target rise above some threshold (Kirchner 1998: 52).

The effect of applying the LAZY constraint to a plosive is that insufficient effort is expended to effect a stop contact. The resulting sound is necessarily, then, some kind of continuant. On the basis of a sample of thirty languages, Kirchner observes (1998: 99, 102) that such a continuant sound, weakly fricated, that results from the lenition of a stop is always a non-strident fricative, of the kind known as a spirant in traditional philology. (For this reason stop lenition is also known as 'spirantization'.) Kirchner explains lenition in these terms:

> Spirantization, under the effort-based approach, is viewed as reduction of the magnitude of a stop gesture, for reasons of effort minimization, to the point where closure is lost. The interval of close constriction in such a reduced gesture is brief, hence the duration of the friction will be correspondingly brief; and because there is no opportunity for a significant build-up of air pressure behind the constriction, the friction will be relatively weak, or completely absent. (Kirchner 1998: 110)

In comparison with a spirant, a strident fricative requires 'a relatively precise, sustained close constriction, in order to generate highly turbulent airflow' (p.111). Kirchner derives the following scale (5) of effort for obstruent articulations (p.118):

(5) Obstruent effort scale:
Strident affricate > strident fricative > stop > non-strident fricative

The lenition of a stop results in a non-strident fricative despite the fact that the non-strident fricative is in itself a marked sound type within phoneme inventories, with respect to a strident fricative. A strident sound, with intense high-frequency noise, less marked in inventories, is better suited to perception. The effort-based theory, then, distinguishes articulatory markedness (which reflects the principle of minimization of effort, and favours non-strident continuants) from perceptual markedness, which favours stridents. The effort scale in (5) is thus distinct from the sonority hierarchy, which tends to reflect perceptibility.

I turn now to examine the contexts which favour or hinder lenition. Lenition is favoured by the openness of adjacent segments, by rapid speech rate, and by a non-careful speech style. The first of these is considered more closely here. Lenition is particularly favoured next to open vowels, between two of which the degree of displacement of the articulatory organs to make a stop contact is great. Kirchner (1998: 182) observed intervocalic lenition in thirty-three languages. Adjacent close vowels favour lenition less than open ones, and adjacent liquids do so less than vowels. Kirchner establishes a hierarchy (6) of adjacent segments provoking lenition according to the degree of jaw opening.

(6) Hierarchy of lenition-favouring contexts (Kirchner 1998: 197):
Low vowel > mid vowel > high vowel > liquid > glide > nasal > stop > strident fricative >...> full or partial geminate

Kirchner's hierarchy (6) of lenition-provoking segments is not entirely adequate for Catalan, as is shown below. In fact, it will be necessary to consider factors other than the degree of jaw aperture. However, from any phonetically motivated hierarchy of lenition-provoking contexts it is possible to derive an inherent ranking of LAZY constraints. Languages will differ in the manner in which they interleave constraints of the LAZY family with faithfulness constraints. Thus a fragment of the constraint hierarchy for the typical pattern of lenition in Catalan is as given in (7). Here faithfulness (IDENT(cont)) comes between the part of the LAZY series that has continuants in the left context and the part of the series in which the left context has consonants with less aperture than continuants. This part, lying below faithfulness, is not active.

(7) $\text{LAZY}_{\text{V_V}}$ » $\text{LAZY}_{\text{r_V}}$ » $\text{LAZY}_{\text{fricative_V}}$ » IDENT(cont) » $\text{LAZY}_{\text{stop_V}}$

In (7) the right-hand context remains the same /_V in all the members of the LAZY series. Without more precise quantification of effort, it is more difficult to derive the universal constraint hierarchy when the right-hand contexts also vary.

For example, does the context between vowel and liquid (/V_liquid) require more or less effort to maintain stop closure than the context between liquid and vowel (/liquid_V)? It follows from Kirchner's theory that the preceding context should weigh more heavily than the following context, inasmuch as upward displacement of the articulatory organ towards stop contact typically counteracts the force of gravity, while displacement away from stop contact is aided by gravity. This aspect is not developed further here; such an approach resembles in outline, if not in detail, that proposed also by Piñeros (2001: 18). This is taken up and developed in §10.1.6, where the Catalan lenition context is considered in more detail.

To summarize, the central idea of the effort-based theory of lenition is that the articulation of each sound, in each context, requires a certain quantity of effort for a perfect realization. The grammars of languages establish thresholds or targets for the amount of effort required in each context (thresholds formalized as constraints of the LAZY family). When the value of this threshold is less than the value required by proper ('faithful') realization of the sound in question, the effects of LAZY appear as lenition.

10.1.4 *Lenition and the enhancement of contrast*

In the preceding section I introduced elements of a functional theory which seeks to explain lenition processes as responding to a principle of minimization of effort. Such a principle reflects an advantage to the speaker as articulator, but its effects are possibly disadvantageous to the hearer, given that non-strident fricatives are somewhat harder to distinguish perceptually than stops. They lack a release burst which contributes to the identification of place of articulation. Alongside functional theories which consider articulatory effort there are also functional explanations which emphasize the goal of facilitating perception. And although lenition may, as just mentioned, undermine the perceptibility of cues to place of articulation, there are reasons to think that lenition may actually facilitate the perception of voicing contrasts. In a study that develops some aspects of Kirchner's approach, Piñeros (2001) draws attention to articulatory problems in the realization of voicing contrast between stops of the same place. Closure of the vocal tract impedes the intended vibration of the vocal folds, inasmuch as vocal fold vibration requires lower air pressure in the supraglottal tract than in the subglottal tract, while oral closure gives rise to an increase in supraglottal pressure. In addition there is an automatic tendency to passive voicing (unintended) of an intervocalic voiceless stop—a tendency which may result in the voicing lenition of voiceless intervocalic stops (Kirchner 1998: 55). These natural articulatory tendencies among oral stops have the result that a voiceless stop is unmarked in initial or final position, and a voiced stop is unmarked in intervocalic position, from a perceptual point of view. Thus, for example, lexical forms /tapat/ and /dabad/ are inclined to merge perceptually as [tabat]. There are several possible articulatory strategies to counter this tendency to neutralize voicing contrasts. But an alternative is to supplement the voicing contrast with

additional effects. For example, one may manipulate the temporal organization of voicing, as happens in English. In English prevocalic voiceless stops VOT is delayed, making the voiceless stops 'aspirated'. At the same time, in English, vowels and nasals are lengthened before voiced stops, an effect which tends to shift the balance of the voicing cues from the stop itself to the preceding vowel or nasal.

Piñeros argues that in Palenquero (a Spanish-based creole spoken in Colombia) the perceptibility of the voicing contrast in stops is enhanced by realizing post-vocalic voiced stops as 'spirants' (non-strident fricatives). In non-strident fricatives, voicing contrasts are effectively realized perceptually, though with some negative consequences for perceptibility of place of articulation. The situation in Catalan is parallel in interesting respects. Moreover, this approach derived from Piñeros's work helps to explain, in a way that Kirchner's LAZY account alone does not, why Catalan lenition affects onset voiced obstruent stops but not coda stops. In Catalan there is never any voicing contrast realized in a coda, either of stops or fricatives. In Optimality terms, the undominated constraint AGREEVoice, together with IDENTOnsetVoice » *VOICEDOBSTR (§5.3), means that LAZY could never work to favour voicing contrast in a coda. (There are other reasons too why there is no constraint having the LAZY effect in codas.) One may suggest, therefore, that the difference between Catalan and Spanish as regards coda lenition (Cat. *a*[bd]*icar* Sp. *a*[βð]*icar*) results precisely from this fact: Spanish maintains a phonemic voicing contrast in codas (*pez* [peθ] 'fish' vs. *red* [reð] 'net', *feliz* [feliθ] 'happy' vs. *Madrid* [maðrið]), while Catalan does not.

10.1.5 *Variation in Catalan lenition*

As mentioned previously (§10.1.1), the tradition in Catalan phonological studies has established a separation between contexts in which obstruent plosives appear (8), and contexts in which voiced fricatives appear (all other contexts than (8)).

(8) Contexts in which voiced stops [b, d, g] appear in Catalan
  i. in codas: *a*[b]*dicar* 'abdicate.INF', *do*[b.b]*le* 'double', *equi*[d]*na* 'echidna', *fle*[g]*ma* 'phlegm'; *ca*[b] *raó* 'no reason', *po*[d] *guarir* 'can cure', *lla*[g] *gelat* 'frozen lake'
  ii. in an onset in the following cases:
    (a) utterance initial: [b]*usca'l!* 'look-for.IMP-DO.3SG.M', [d]*igues!* 'say.2SG.IMP', [g]*ol!* 'Goal!'
    (b) after a [−cont] segment: *cam*[b]*ra* 'room', *ven*[d]*a* 'sale', *tin*[g]*uem* 'have.1PL.PRS.SUBJ'; *ab*[d]*icar* 'abdicate.INF', *pot* [b]*lau* 'blue pot', *un xic* [d]*ifícil* 'a bit difficult', *drap* [g]*astat* 'worn-out cloth'
    (c) after a lateral (only dental stop): *vol*[d]*rà* 'want.3SG.FUT', *gall* [d]*indi* 'turkey'
    (d) after a labio-dental (only bilabial stop): *xef* [b]*o* 'good chef'

In all the contexts other than those in (8), i.e. in onset position, following a vowel or following consonants not specified in (8), it is implied in the literature, if it is not stated explicitly, that the continuant allophones [β, ð, ɣ] are found categorically. It is mostly the dialect studies which present a more complex picture. Recasens (1991b: 182–250) gives a great deal of information about the nature of diatopic variation in lenition, but he gives no quantitative information, and it is not easy to derive an overall picture.

The publication in 1998 of the first volume of the so-called 'ethnotexts' of the Catalan Dialect Atlas (*Atles Lingüístic del Domini Català*) allows one to explore more fully the issue of variability in lenition. Veny & Pons (1998) have edited an anthology, in phonetic transcription, of monologues that were recorded as part of the dialect research programme associated with the *Atles Lingüístic del Domini Català*, now in course of publication. The 1998 volume contains only the monologues recorded in the eastern dialect bloc. The texts are of widely varying length. There are 118, quite evenly distributed geographically, from the territory of continental eastern Catalan, and a further thirty-two from the Balearic islands. On the continent and in Majorca the recordings were made between 1964 and 1977, while those in Minorca and Ibiza were recorded in 1993 and 1994. This time difference needs to be remembered when it comes to interpreting the data derived from the recordings.

I have counted, in the transcribed texts, the stop and fricative variants of the non-strident obstruent phonemes /b, d, g/ as they occur in the contexts listed in (9). The transcriptions confirm that the stop variants are indeed categorical in the contexts set out in (8) above.

(9) Contexts in which lenition variants appear
1. Word-initial following a rhotic: /r ∼ ɾ#__ (e.g. *hort de secà* 'irrigated smallholding')
2. Medial following a rhotic: /r ∼ ɾ.__ (e.g. *garbell* 'sieve')
3. Word-initial following a lateral: /l#__, /ʎ#__ (e.g. *el gos* 'the dog')
4. Medial following a lateral: /l.__ (e.g. *alguna* 'some.F')
5. Word-initial following a sibilant: /z#__, /ʒ#__ (including cases of d͡z#__, d͡ʒ#__) (e.g. *uns ganivets* 'some knives')
6. Word-initial following a vowel: /V#__ (e.g. *la gruixuda* 'the thick-one.F')
7. Medial following a vowel: /V.__ (e.g. *bleda* 'beet') (In this context the number of stop variants is negligible, in continental eastern Catalan.)

The data for continental eastern Catalan and that for the Balearic Islands have been counted separately, in accordance with the hypothesis mentioned in earlier studies that the presence of a voiced labio-dental phoneme /v/, in Balearic, may affect the realization of /b/ in particular.[2] I have not introduced further geograph-

[2] Only one of the continental texts (Alió) comes from a variety that maintains the phoneme /v/. The text is very short in any case, and is excluded from consideration here.

ical divisions within the eastern continental Catalan data, so as to retain an overall perspective. (And in fact systematic differences between regions of eastern continental Catalan have not been observed.)

The raw figures and percentages of voiced bilabial, dental, and velar fricative realizations in the contexts listed in (9) are displayed in Table 10.1.

TABLE 10.1 Percentage of voiced fricatives in six phonological contexts in the eastern continental Catalan texts of Veny & Pons (1998)

Context	No. of cases of /b/	% [β]	No. of cases of /d/	% [ð]	No. of cases of /g/	% [ɣ]
1. /r ~ ɾ#__	10	40.00	22	13.64	–	–
2. /r ~ ɾ.__	64	87.50	58	84.48	23	86.96
3. /l#__, /ʎ#__	169	42.60			36	36.11
4. /l.__	6	33.33			24	75.00
5. /z#__, /ʒ#__	172	84.30	310	85.48	73	93.15
6. /V#__	670	92.09	771	93.00	113	92.04
Total	1,091	82.13			269	82.90
Total excluding contexts 3 and 4	916	89.74	1,161	89.06	209	91.86

The figures in Table 10.1 demonstrate that between the categorical realization of stops in certain contexts (those of (8)) and the categorical realization of fricatives in others ((9) 7, medial post-vocalic) there is a set of contexts (namely (9) 1–6) in which variation is to be found, at least among the speakers interviewed for the Atlas. Considering first the role of place of articulation, essentially, in each of the six contexts there are no important differences in respect of place of articulation of the non-strident obstruent. This is more evident in the cells where there is a large number of cases. For example, in context 2 (medial following rhotic), the presence of fricative realizations ranges only between 84.48 per cent for the dental ([ð]) and 87.50 per cent for the bilabial ([β]). In context 6 (word-initial post-vocalic) the occurrence of fricative variants ranges only between 92 and 93 per cent. Excluding the post-lateral context in which the behaviour of the dental obstruent (categorical stop, as in (8)) is very different from that of the bilabial or the velar, the proportions of fricative realizations (final row of Table 10.1) are: bilabial [β] 89.74 per cent, dental [ð] 89.06 per cent, and velar [ɣ] 91.86 per cent. That is to say, the proportion of fricatives at each of the three places of articulation is virtually identical.

Secondly, it is clear that, when there are sufficient data to distinguish, word-initial position favours fricative realization somewhat less than medial position.[3]

[3] The figure in only one cell is out of line: in contexts 3 and 4 the figure for medial [-1.β-] (33.33%) is lower than that for initial [-1#β-] (42.60%). Given the very low number of cases of /-1.β-/, I do not regard this as significant.

The most likely interpretation of this difference is that output–output correspondence brings to bear the analogical model of the words' citation forms (utterance-initial, with stops). Positional faithfulness considerations (Beckman 1998) might also be expected to contribute to the difference between initial and medial realizations. That is, the psychological salience of a strong position (word-initial in this case) would oppose some degree of resistance to weakening or 'LAZY' tendencies.

On the basis of the data in Table 10.1 one can establish a hierarchy (10) of contexts favouring the (variable) fricative realization of voiced non-strident obstruents in eastern continental Catalan. In the hierarchy in (10) the numerical data are simplified, taking the means of the three places of articulation and rounding the figures. Below the medial intervocalic position (virtually 100 per cent) omitted from Table 10.1, comes the word-initial postvocalic position (93 per cent fricatives). Tied in second place are the medial position following a rhotic and the initial position following a sibilant (86 per cent fricatives). In third place comes the medial post-lateral position for labials and velars (67 per cent fricatives). In fourth place comes the word-initial post-lateral position for labials and velars (41 per cent). And in last place comes word-initial position following a rhotic (22 per cent).

(10) /V#__ > /r ~ ɾ.__; /z#__, /ʒ#__ > /l.__ > /l#__, /ʎ#__ > /r ~ ɾ#__
93 86 67 41 22

The data for the Balearic dialects are set out in Table 10.2, with separate subtables for the transcribed data from each of the three islands, as there are marked differences between the data for Majorca and those for Minorca and Ibiza (see the bilabial and dental columns). The rows for the totals excluding context 7 (medial post-vocalic) provide the appropriate comparison with the figures of Table 10.1 for eastern continental Catalan, where the medial post-vocalic context was not included.

It is not possible to be sure whether the differences between Majorca and the other two islands in the proportions of fricative realizations in Table 10.2 reflect diatopic or diachronic differences, given that almost thirty years elapsed between the two periods of data collection. What is evident, however, is a notable difference between the data of the islands taken as a whole and those of continental eastern Catalan given in Table 10.1. The first difference to comment on is a marked difference in the Balearic data between the obstruents of different places of articulation. Whereas in continental eastern Catalan the proportion of fricatives of all places in comparable contexts ranges between 89 and 92 per cent, in Balearic taken as a whole the corresponding figures range from 82.05 per cent for the velar [ɣ], through 69.08 per cent for the dental [ð], coming down to 29.67 per cent for the bilabial [β]. There can be little doubt that the much greater frequency of the bilabial plosive [b] in lenitable contexts in Balearic

TABLE 10.2 Percentage of voiced fricatives in seven contexts in the texts in Veny & Pons (1998) from Majorca, Minorca, and Ibiza. Rows with too few examples have been omitted.

Majorca (1964-65)	No. of cases of /b/	% [β]	No. of cases of /d/	% [ð]	No.of cases of /g/	% [ɣ]
1. /r ~ ɾ#__	9	44.44	14	7.14	–	–
2. /r ~ ɾ.__	15	13.33	20	30.00	–	–
5. /z#__, /ʒ#__	15	13.33	41	34.15	11	81.81
6. /V#__	39	15.38	234	54.70	27	74.07
7. /V.__	75	32.00	145	87.59	103	98.06
Total	153	24.84	454	60.79	141	92.20
Total contexts 1-6	78	17.95	309	48.22	38	76.32
Minorca (1994)						
1/2. /r ~ ɾ(#)__	–	–	7	28.57	–	–
5. /z#__, /ʒ#__	14	57.14	48	93.75	–	–
6. /V#__	25	28.00	95	91.58	12	91.67
7. /V.__	30	66.67	48	87.50	64	100.00
Total	69	50.72	198	88.89	76	98.68
Total contexts 1-6	39	38.46	150	89.33	12	91.67
Ibiza (1993)						
1/2. /r ~ ɾ(#)__	–	–	12	66.67	–	–
3/4. /l(#)__, /ʎ#__	9	11.11			10	60.00
5. /z#__, /ʒ#__	13	46.15	77	83.12	7	100.00
6. /V#__	43	41.86	128	87.50	11	100.00
7. /V.__	38	73.68	52	100.00	47	100.00
Total	103	51.46	269	87.73	75	94.67
Total contexts 1-6	65	38.46	217	84.79	28	85.71

reflects a principle of maintenance of contrast between this phoneme and the voiced labio-dental fricative /v/, a phoneme which is lacking in the eastern continental variety. Observe that even in the context most favourable to lenition, namely, medial post-vocalic position, the proportion of stop realizations does not fall below 26 per cent (Ibiza); in the majority of contexts it remains above 50 per cent.

The figures presented for Balearic in Table 10.2, taken as a whole, seem to indicate a weaker preference for fricative realizations than on the continent, which might reflect a general phonological conservatism, for which there is some other evidence (including maintenance of /v/, just mentioned, and maintenance of stressed /ə/). Nevertheless, the low proportion of dental fricative realizations [ð] in Balearic is very largely due to the figure from Majorca (48.22 per cent). In Majorcan the dental fricative is clearly the minority variant

in all the post-consonantal contexts. The mean figure for the dental fricative in the other two islands, 86.65 per cent in the matching contexts, is hardly less than the figure for eastern continental Catalan (89.06 per cent). Similarly, it is Majorcan that provides a significant proportion of velar stops in lenitable contexts, 23.68 per cent in contexts 5 and 6. It should be borne in mind too that the Majorcan texts lack examples of velar obstruents following laterals, due largely to the fact that there the masculine definite article ends not with a lateral (*el*) but with a sibilant (*es*). In continental eastern Catalan the post-lateral context is the one that provides the lowest proportion of velar fricatives (51.66 per cent). Thus if this post-lateral context were as frequent in Balearic as it is on the mainland, there would be an opportunity for the proportion of velar stops in Majorca to be even higher than it is.

To summarize, the distribution of variation between voiced stops and fricatives differs between the islands and the mainland in three respects. On all three islands the frequency of bilabial stop variants in lenitable contexts is much greater, a fact which is plausibly due to the need to distinguish the bilabial from the labio-dental obstruent in manner of articulation and not just in place. The low proportion of fricatives at all three places of articulation is characteristic of Majorcan alone. In this variety only the velar in medial post-vocalic position reaches a level of fricative realizations comparable to the norm in the eastern continental Catalan texts. It is not implausible, then, that the lower proportion of fricatives in general in Majorcan reflects phonological conservatism. The more modern tendency would be what we see in the continental texts roughly contemporaneous with the Majorcan ones, and in the texts of a generation later from Minorca and Ibiza. Further research is needed to determine whether the differences between Majorca and the other islands are to be interpreted better as diatopic or diachronic (or both).

One can observe in the data from all the islands a hierarchy of lenition based on place of articulation, with the bilabial /b/ more resistant to lenition than the dental /d/, which is in turn more resistant than the velar /g/. Such a hierarchy stands out more if we exclude from the picture the data from medial post-vocalic position where fricative realizations approach categoricity (a context which essentially reaches categoricity in the continental texts). This place hierarchy for voiced stops, /b/ > /d/ > /g/, matches what is found in other languages (Croft 2003: 158–60), and reflects degrees of aerodynamic difficulty in the articulation of a voiced stop. During the articulation of a voiced stop, the smaller the volume of air contained between the glottis and the place of articulation the harder it is to maintain the pressure difference between the subglottal and supraglottal air-stream that is necessary in order for the vocal folds to vibrate.

The hierarchy of lenition-favouring contexts is slightly different in Balearic varieties from what is observed in eastern continental Catalan. Consider first the bilabial obstruent. Excluding the context following a lateral, where we only have data from Ibiza, the following hierarchy can be observed (11): the most favourable context for fricative realization is word-initial following a rhotic (data only

from Majorca). Next comes medial post-vocalic position, followed by post-sibilant position and medial post-rhotic position. In the last position comes word-initial post-vocalic position. In (11) the figures are the rounded percentages of bilabial fricatives in each context.

(11)	/r ~ ɾ#__	>	/V.__	>	/r ~ ɾ.__	>	/z#__, /ʒ#__	>	/V#__
Majorca	44		32		13		13	≈	15
Minorca			67				57		28
Ibiza			74				46		42

Observe that the hierarchy in (11) is the reverse of that seen in (10). If it were only articulatory effort that contributed to lenition it would be difficult to make sense of such a difference as exemplifying typological variation between related speech varieties. However, I believe the reversed hierarchy of contexts favouring bilabial stop lenition in Balearic supports the interpretation that attributes bilabial stop preservation to the enhancement of phonemic contrast with the labio-dental fricative phoneme. On this basis one might expect phonemic contrast to be better preserved in word-initial position, reflecting the important role of the initial phoneme for word recognition (Beckman 1998). A functional explanation is harder to find for the position at the top of the hierarchy in (11) of the word-initial post-rhotic context.[4]

Turning back to the voiced dental and velar obstruents, the hierarchy of contexts favouring lenition in Balearic is again essentially the same as that seen in eastern continental Catalan (10). The Balearic version is set out in (12), again with rounded percentages of fricatives in the contexts listed. Figures that appear out of line are commented on in notes beneath the hierarchy.

(12)		/V.__	>	/V#__	>	/z#__, /ʒ#__	>	/r ~ ɾ.__	>	/r ~ ɾ#__
Majorca	ð	88		55		34		30		7[5]
	ɣ	98		74		*82				
Minorca	ð	†87		92	≈	94		29		
	ɣ	100		92						
Ibiza	ð	100		87		83		67		

*One fewer case of [ɣ] would be sufficient for the figure in this cell to correspond to its position in the hierarchy.

†Two more cases of [ð] would be sufficient for the figure in this cell to correspond to its position in the hierarchy.

[4] The 44% figure represents four cases of [-ɾ#β-], all from the same text, from Santanyí. In all four cases the word-final rhotic realizes an input /z/, and the rhotic is transcribed with a small superscript letter: *dos brins* [doʳ βrins] 'two bits' (two occurrences), *dos brinets* [doʳ βriˈnəts] 'two bits.DIM', *es braç* [əʳ βraz] 'the arm'. Perhaps too much should not be made of these examples.

[5] Of the 14 examples of /-ɾ#d-/ in Majorcan, seven (all with [-ɾ#d-]) are from the same speaker, also from Santanyí, but not the speaker who provided the example of [-ɾ#β-]. The Santanyí texts offer no case of medial /-ɾb-/ or /-ɾd-/. The surprising position of the context /r ~ɾ#__ in this hierarchy may thus be due to a fortuitous bias in the data.

Setting aside the cases, rare in Balearic, of the velar obstruent following a lateral, the hierarchy of contexts favouring a dental or velar fricative in Balearic (12) is in accord with the eastern continental hierarchy (10). The hierarchy in (13) combines the two, giving a ranking of contexts favouring a non-strident fricative in the eastern Catalan dialect bloc. The hierarchy in (13) is interpreted as reflecting two types of influence, in principle independent, as represented in (13a) and (13b).

(13) Medial post-vocalic > word-initial post-vocalic > word-initial post-sibilant > medial post-rhotic > medial post-lateral > word-initial post-lateral > word-initial post-rhotic

(13a) Medial > word-initial

(13b) Post-vocalic > post-sibilant > post-lateral, post-rhotic (> after [−cont])

As mentioned above, the hierarchy in (13a) reflects the principle of Positional faithfulness, in favouring the preservation of input features in a psychologically salient position. Or what comes to much the same thing: preservation of initial stops would reflect Word-Phrase correspondence constraints, given the realization of word-initial voiced non-strident obstruents as stops in utterance-initial or citation forms. Hierarchy (13b) would be the one corresponding to phonetic contexts conditioning the expenditure of effort.

10.1.6 *Further interpretation of the hierarchy of leniting contexts in Catalan*

So far in section §10.1 the context following the consonant liable to lenition has not been examined. I have merely stated that lenition affects consonants in onsets but not in codas. Recall that an obstruent in onset position must be followed (a) by a vowel (e.g. /ɡa/, or (b) by a semivowel (e.g. /ɡwa/) or a liquid (e.g. /ɡla/). All these following contexts display a sufficient degree of jaw aperture to favour lenition, in accordance with Kirchner's approach (see (6) above). In principle, therefore, one might expect to find lenition more favoured before vowels than before liquids, given that we have seen vowels and liquids in the preceding context contributing differently to proportions of lenition variants. The transcribed texts examined here provide no support for this expectation. Although Kirchner does not mention it, it would be legitimate to argue, as in fact Piñeros does (2001: 17–19), that the nature of the preceding segment has a stronger influence on the expenditure of effort than the following segment. Articulatory displacement from a vowel to a stop, requiring effort, is more important than subsequent displacement, which does not affect directly whether consonant contact is realized or not. The nature of a following segment may provoke an anticipatory effect so as to minimize effort to come, but that is in principle an independent effect (which may be added, naturally, to that of LAZY which directly affects the realization of a stop). It is not surprising, then, that there are languages in which the lenition context is determined exclusively in terms of

the preceding segment. Kirchner mentions several, to which Spanish can be added.

There is an important respect in which the Catalan lenition context hierarchy (13b) diverges from that proposed by Kirchner, given in (6) above, and reproduced here as (14).

(14) Low vowel > mid vowel > high vowel > liquid > glide > nasal > stop > strident fricative >. . . > full or partial geminate

The important difference between (13b) and (14) lies in the place of the strident fricative in the hierarchy. While according to Kirchner's account adjacent strident fricatives favour lenition less than nasals or stops, (13) has shown that in Catalan sibilants (which are strident fricatives) favour lenition more than liquids, and evidently more than nasals or stops which block lenition completely. It appears that for lenition to take place in Catalan, as in the other languages of Iberia, it is absolutely required that the preceding segment be a continuant. The defining characteristic of continuant ([+cont]) segments is that they are produced with a continuous airflow through the oral tract. I suggest that the airflow of a continuant vowel or consonant is a force which is opposed to the realization of an oral stop. To realize a stop contact, the articulator not only has to be displaced but also has to resist the opposing force of a current of air. It is this current of air that tends to impede the realization of an oral stop. By contrast, when there is no preceding airflow, the oral cavity does not offer resistance to the formation of a stop contact. For this reason a stop is relatively unmarked in utterance-initial position, as it is also following a nasal consonant, or following another stop. In such contexts—for the languages of Iberia at least—differences of jaw aperture are never sufficient to allow lenition. To take another case, a preceding lateral, articulated with full contact of the centre of the front of the tongue ([+coronal]) behaves exactly like a preceding stop with respect to a following coronal stop, inasmuch as there is no need for the tongue to abandon the stop contact already existing. Indeed, all that is necessary is that it should not do so. To pass from a lateral to a non-lateral, or from alveolo-palatal to dental articulation, some other modificatory gestures are needed, but these do not prejudice the maintenance of stop contact by the front of the tongue.

Among the continuant consonants—sibilants, laterals, rhotics—that favour lenition there are differences in manner of articulation that might tend to promote or inhibit lenition, such as to result in the hierarchy (13b). A sibilant requires a strong airflow, in order to produce turbulent noise; it is for this reason, I believe, despite the reduced degree of jaw aperture, that sibilants come next to vowels in the Catalan lenition context hierarchy. In contrast, in both laterals and rhotics there is some kind of obstruction of the airflow that reduces the force opposed to the realization of stop contact even when made at a different point of articulation. The contact of a tap, or each of the series of contacts of a trill, is a form of stop tending to reduce the flow of air. It has been proposed that a tap is a contour segment with three phases [[+cont][−cont][+cont]]. The non-continuant

element, despite its brevity, would be enough for a rhotic to act in a similar way to a nasal or a stop, though to a much lesser degree. Because of some partial hindrance of the airflow in both laterals and rhotics, therefore, the context following a liquid is less favourable to fricativization than the position after a vowel, and even to the position after a sibilant.

10.1.7 *Remaining issues*

In §§10.1.2–10.1.6 I gave attention to some of the issues in Catalan lenition set out in (1). In this section I give some attention to the rest. As far as issue (1. ii) is concerned, it will be evident that I take the view that phonetics can make a major contribution to explaining the phenomenon of lenition. In fact, I do not support a contrast or opposition between phonetic and phonological explanations of the type discussed by Palmada (1997). While there may be processes, or alternations, that are not sufficiently explained by synchronic phonetics (or synchronic morphology, for that matter), it now seems odd to argue that phonetic and phonological theory are in conflict and compete to answer the same questions, such that science can declare one the winner.

The similarities or parallels between the lenition discussed here and alternations between voiced alveolo-palatal affricate and fricative mentioned in (1. iv; *fu*[t͡ʃ] – *fu*[ʒ]*im*) are, I believe, rather superficial. The most obvious parallel feature is the presence of a stop (affricate in the alveolo-palatal case) in a coda, alternating with a medial intervocalic fricative. But the voiced alveolo-palatal fricative is not a realization typical of all of Catalan in the way that voiced intervocalic non-strident fricatives are, and where the voiced alveolo-palatal fricative exists it participates in a series of contrasts very different from those shared by the non-strident fricatives (§2.1.3.3). The voiced alveolo-palatal fricative is opposed to an intervocalic voiced geminate affricate (*pruïja* [pɾuˈiʒə] 'itch' – *pitja* [ˈpid͡dʒə] 'tread.3SG.PRS.IND'), to a voiceless fricative (*pruïja* – *pixa* [ˈpiʃə] 'piss.3SG.PRS.IND'), and to a voiceless affricate at the same place of articulation (*pruïja* – *fitxa* [ˈfit͡ʃə] 'file'). Moreover, the distribution of voiced alveolo-palatal fricatives and affricates in word-initial position and following a consonant, even when there is no contrast between them, is very different, as regards the linguistic, sociolinguistic, and geolinguistic contexts, from what is observed with the non-strident obstruent fricatives and stops (§2.1.3.3). The most one can say is that the phenomena have some aspects in common, which is very far from saying that they result from the same causes in the same conditions.

It is certainly legitimate to call lenition an assimilation process (see (1. vii)), but this label, established in Generative Phonology, contributes little to understanding the process, inasmuch as it does not help to specify its effect, its context, or its cause. One advances further by identifying in lenition, as I have attempted above, a weakening process related to natural aspects of the articulation of voiced stops, and at the same time a mechanism that serves to enhance voicing contrast

in contexts where such contrasts is endangered by problems inherent in the articulation of voiced and voiceless stops.

The matter of underlying or lexical representation (1. viii) is important inasmuch as it is right not to take it for granted. It is true that, in a case of categorical complementary distribution of two (or more) allophones, Optimality Theory constraints should govern directly and explicitly the appearance in the output of both allophones equally, in accordance with the theories of Richness of the Base and lexicon optimization (McCarthy 2002: 66–70, 76–82). In such a case, the choice of underlying representation is not significant. (For this reason Kirchner (1998: 95) argues that constraints with lenition effects should have complementary constraints in the grammar enforcing fortition in all the complementary contexts.) But the distribution of stops and fricatives in Catalan is not perfectly complementary and categorical, at least in those varieties studied in §10.1. What we find is variable stops in positions highly favourable to lenition, while we do not find fricatives, even variably, in putative fortition contexts. Such distribution implies that the lexical representation really does have voiced stops (and that is, of course, also the historical situation). But this situation may not continue, and may not even be appropriate for all Catalan varieties. Richness of the Base and lexicon optimization oblige the researcher to justify an input representation if it is not indeterminate.

10.2 CONSONANT/ZERO ALTERNATIONS

10.2.1 *N/zero alternation*

A salient feature of Catalan morphophonemics involves a large number of stem morphemes (roots or derivational suffixes) which are vowel-final when no suffix follows but which show [n] before a suffix, whether inflectional or derivational. Some examples among nominals are given in (15), taken from Bibiloni (2002: 273). The orthography indicates the alternation in a straightforward manner.

(15)

Base	Plural /+z/	Feminine.SG /+a/	Derived form
pa 'loaf'	*pans*		*paner* 'bread basket'
català 'Catalan'	*catalans*	*catalana*	*catalanista* 'Catalanist'
amè 'agreeable'	*amens*	*amena*	*amenitat* 'agreeableness'
fi 'fine'	*fins*	*fina*	*finesa* 'fineness'
menorquí 'Minorcan'	*menorquins*	*menorquina*	*menorquinitat* 'Minorcan-ness'
so 'sound'	*sons*		*sonar* 'sound.INF'
lliçó 'lesson'	*lliçons*		*alliçonar* 'instruct.INF'
comú 'common'	*comuns*	*comuna*	*comunitat* 'community'

The pattern of alternation in (15) is extraordinarily frequent in the lexicon, not least because some very productive suffixes display the alternation, such as the adjective-forming suffix *-í* ∼ *-ina*, etc., as in *menorquí* 'Minorcan', or the noun-forming suffix *-ió*, as in *fusió* 'fusion', cf. *fusionar* 'to fuse'. In Generative Phonology, from Lleó (1970: 7–8) to Bonet & Lloret (1998: 100), these alternations were accounted for by a rule of word-final *n*-deletion. But the rule was always somewhat problematic (see Wheeler 1975/9: 270–75). There are lexical exceptions to it (see below). Though *n*-deletion is frequent in nominals, only two verb roots display deletion (*venir* 'come' and *tenir* 'have'), in contrast to some fifteen or so verb roots which lack *n*-deletion in the relevant contexts. In the vast majority of stems with *n*/zero alternation, the phenomenon appears, as in the examples in (15), after a stressed vowel. In (16) I give a complete list of the roots illustrating *n*/zero alternation after an unstressed vowel.

(16)

Stem	Plural /+z/	F.SG /+a/	derived form
argue 'windlass'	*àrguens* 'frame for pannier'		*arguenell* 'cat-head'
			arguenells 'saddlebag'
ase 'ass'	*ases* ∼ *àsens*		*asenada* 'stupidity'
cove 'conical basket'	*coves* ∼ *còvens*		*covenada* 'basketful'
	créixens 'cress'		*creixenar* 'cress bed'
freixe 'ash tree'	*freixes* ∼ *fréixens*		*freixeneda* 'ash plantation'
home 'man'	*homes* ∼ *hòmens*		*homenia* 'virility'
jove 'young'	*joves* ∼ *jóvens*		*jovenalla* 'young people'
marge 'bank'	*marges* ∼ *màrgens*		*margener* 'bank-builder'
orfe 'orphan.M'	*orfes* ∼ *òrfens*	*òrfena*	*orfenesa* 'orphanhood'
orgue 'organ'	*orgues* ∼ *òrguens*		*orguener* 'organ-builder'
rave 'radish'	*raves* ∼ *ràvens*		*ravenar* 'radish field'
terme 'limit'	*termes* ∼ *térmens*		*termenar* 'set.INF-limits'
verge 'virgin'	*verges* ∼ *vèrgens*		

In the case of the type (16) examples, there are alternative plural forms. The plural forms in *-ns* are characteristic of western Catalan dialects; those without [n] are characteristic of eastern dialects. From an OT perspective it might appear

that there are constraints such as *Vn]$_{Wd}$, *V̌ns]$_{Wd}$. The fact that verb forms resist *n*-deletion might be an Optimal Paradigms effect: verbs have over fifty inflected forms in which a stem-final consonant is not word-final, and these would exert strong paradigm pressure against stem allomorphy. And *tenir* and *venir* might display a weaker OP effect than other verbs because, alongside /ten/ and /ben/ root allomorphs, these two verbs also have /tin/ and /bin/ allomorphs respectively in a number of contexts. Final [n] in third person plural forms (*roguen* [ˈrɔɣən] 'ask.3PL.PRS.IND', *sabien* [səˈβiən] 'know.3PL.PST.IMPF') might be protected by a REALIZEMorpheme constraint.

The major difficulty with treating *n*-deletion as a true phonological process comes from the substantial number, if not large, proportionately to the total with *n*-deletion, of lexical exceptions among nominals. In (17) I list some examples of content nominals lacking *n*-deletion after a stressed vowel.

(17)

Aran (toponym)	*mossèn* 'reverend'
ban 'edict'	*nan* 'dwarf'
Camprodon (toponym)	*nen* 'child'
colon 'colonist'	*nin* 'doll'
con 'cone'	*pregon* 'deep'
escon 'bench'	*Ramon* (given name)
Ferran (given name)	*saxofon* 'saxophone'
gen 'gene'	*segon* 'second (A)', 'second (N)'
gran 'large'	*son* 'sleep'
Joan (given name)	*tren* 'train'
món 'world'	*tron* 'throne'

In (18) are some function words without *n*-deletion after a stressed vowel.

(18)

algun 'some'	*quan* 'when'
arran 'adjacent to'	*quin* 'which?'
cadascun 'each	*un* 'one, a'
on 'where'	

The examples in (17) and (18) suggest that the lexical difference between *n*-deletion and its absence does not come down to a motivated partition of the vocabulary such as 'native' vs. 'foreign'. In fact, the 'opacity' of *n*-deletion arose historically from a sequence of sound changes: (a) /n/ > Ø /__#, (b) /nd/ > /nn/, (c) /nn/ > /n/, all of which occurred probably at least a thousand years ago. The lexical incidence of the two patterns varies little between dialects (Recasens 1991b: 260); there are a few well-established cases of 'over-application', from an etymological perspective, such as *gla* ∼ *aglà* 'acorn' (< GLANDEM), *segó* 'bran' (< SECUNDUM), *rodó* 'round' (< ROTUNDUM).[6] Examples of final [n] in nominals retained after an unstressed vowel generally come in relatively technical terminology, such as *androgin*

[6] But cf. the expected [-n] in the toponym *Camprodon* in (17) < CAMPUM ROTUNDUM 'round field'.

'androgynous', *dictamen* 'verdict', *examen* 'examination', *fenomen* 'phenomenon', *oxigen* 'oxygen', *monòton* 'monotonous', *pol·len* 'pollen', and so on.

Traditional grammar deals with the pattern in (15) in terms of '*n*-insertion'. For example, the *Diccionari de la Llengua Catalana* (Giralt et al. 1993: 13) gives plurals in *-ns* as the general rule for nominals ending in a stressed vowel. Only the exceptions to this rule are mentioned specifically in the body of the dictionary. Such 'exceptions', then, are words ending in a stressed vowel in the citation form that do not show [n] in any inflected or derived forms. Many of these exceptions are borrowings, but not all are; examples are given in (19).

(19)

cafè 'coffee'	*lilà* 'lilac'	*puré* 'purée'
ce '(letter) C'	*mercè* 'favour'	*quinqué* 'oil lamp'
comitè 'committee'	*nu* 'naked'	*sofà* 'sofa'
cru 'raw'	*oboè* 'oboe'	*tabú* 'taboo'
do '(note) C'	*pagaré* 'IOU'	*vostè* 'you (polite)'
esquí 'ski'	*papà* 'daddy'	*ximpanzé* 'chimpanzee'
fe 'faith'	*peroné* 'fibula'	

Examples like (19) taken together with those of (17) and (18) show that, while *n*/zero alternation is the default pattern for lexemes whose unmarked (base) form ends in a stressed vowel, there is no general *phonological* process of *n*-deletion, or of *n*-insertion. Since *n*-deletion originally took place, as a sound change, in morphological bases alone (singular, and mostly masculine gender), it would not be surprising for 'rule reversal' to have ensued. But there is not a great deal of evidence that it has, or not to the extent of morphological alternation becoming truly rule-governed ('insert /n/ after a stressed final vowel before an affix'). Only a few lexemes are observed as developing unetymological (and non-standard) [n] in plural forms: *nus* [ˈnuns] 'naked.M.PL' *sofàs* [suˈfans] 'sofas', *altars* [əlˈtans] 'altars', *pilars* [piˈlans] 'pillars', *tambors* [təmˈbons] (Bibiloni 2002: 277). Note that the last three of these are actually lexemes with historical final /R/, as the orthography suggests; see §10.2.2.

In classical Generative Phonology all morphophonemics was dealt with by 'phonological rules' insofar as processes were of any generality at all, because there was no separate 'morphological component'. Conflicting patterns were dealt with by 'minor' rules, or by lexical exception features. But such an approach is not appropriate in a phonological theory in which functional motivation is sought for phonological constraints. So a realistic OT account involves the morphemes of the type illustrated in (15) and (16) displaying allomorphy. Each of those morphemes comes in two forms, for example, in the case of *so* 'sound', /ˈsɔ/ and /ˈsɔn/. Then, formally, one might imagine a $*n]_{Wd}$ constraint (20)—perhaps an anti-alignment constraint—ranked below faithfulness to /n/ (IDENT/n/, MAX/n/). Thus, if and only if an allomorph without -n] is available /-V/ is preferred to /-Vn/ by $*n]_{Wd}$. It still needs to be shown why the /-n/ allomorph is preferred in other contexts, despite paradigm uniformity violations. In the context before a vocalic suffix, as in *sonar* /sɔn+ˈa+R/ 'sound.INF', the

allomorph /sɔn/ is preferred by ONSET; [suˈna] is more harmonic than *[suˈa] (vowel reduction being undominated here). Before a consonantal suffix, such as /+z/ 'PLURAL', I suggest as a possibility that /sɔn/ is preferred to /sɔ/ by an alignment constraint that requires stems with final stress to end in a consonant. Such a constraint (21) would in principle promote the matching of stress and syllable weight (= SWP §8.5 (20c)), though its effect in Catalan is normally concealed by faithfulness constraints such as DEPC.

(20) *n]$_{Wd}$: A phonological word does not end in /n/. (An apparently unnatural constraint as so expressed, but actually abbreviating constraints of faithfulness to all consonants other than /n/—IDENTC¬n/MAXC¬n—ranked above *CODA.)

(21) ˈσSTEM-C]: A stem having a stressed final syllable ends in a consonant.

The effects of constraints selecting between allomorphs with and without final /n/ are illustrated in tableau (22), where the examples are (22a) *so* 'sound', plural *sons*, derived verb *sonar* 'to sound'; (22b) *freixe* 'ash tree', plural *freixes*, derived noun *freixeneda* 'ash plantation'; (22c) *nen* 'child.M', plural *nens*, *nena* 'child.F'; and (22d) *mercè* 'favour', plural *mercès*. Varieties in which the plural of *freixe* is *fréixens* display in the place of (21) a more general constraint, namely STEM-C]: A stem ends in a consonant.

(22)

		ONSET	DEPC	IDNAS	MAX/n/	*n]$_{Wd}$	ˈσSTEM-C]	UNIF	PU
(a) /ˈsɔ/~ /ˈsɔn/									
i. /ˈsɔ/	☞ ˈsɔ						*		
/ˈsɔn/	ˈsɔn					*!			
ii. /ˈsɔ+z/	ˈsɔs						*!		
/ˈsɔn+z/	☞ ˈsɔns								*
iii. /ˈsɔ+ˈa+R/	suˈar	*!					*		
/ˈsɔn+ˈa+R/	☞ suˈnar								*
(b) /ˈfreʃe/ ~ /ˈfreʃen/									
i. /ˈfreʃe/	☞ ˈfreʃə								
/ˈfreʃen/	ˈfreʃən					*!			
ii. /ˈfreʃe+z/	☞ ˈfreʃəs								
/ˈfreʃen+z/	ˈfreʃəns								*!
iii. /ˈfreʃe+ˈɛd+a /	freʃəˈɛðə	*!							
/ˈfreʃen+ˈɛd+a/	☞ freʃəˈnɛðə								*

(continued)

			ONSET	DEPC	IDNas	MAX/n/	*n]$_{Wd}$	'σSTEM-C]	UNIF	PU
(c)	i. /'nɛ$_1$n$_2$/									
		'nɛ$_1$				*!		*		
		'nɛ$_{12}$			*!			*	*	
		☞ 'nɛn					*			
	ii. /'nɛn+z/	'nɛ$_1$s				*!		*		*
		☞ 'nɛns								
	iii. /'nɛn+a/	'nɛə	*!			*		*		*
		☞ 'nɛnə								
(d)	i. /mer'sɛ/	☞ mər'sɛ						*		
		mər'sɛn		*!						
	ii. /mer'sɛ+z/	☞ mər'sɛs						*		
		mər'sɛns		*!						*

ONSET, DEPC, IDENTNas, MAX/n/ » *n]$_{Wd}$ » STEM-C] » UNIFORMITY, **PU**

It is perhaps of some interest to reconstruct the sequence of constraint reranking by which the situation illustrated in (22) was reached historically. I assume, abstracting away from other considerations, that the original alternation for *so* 'sound' involved the forms and ranking in (23a).

(23) (a) [ˈsɔn] – [ˈsɔns] – [soˈnaɾ]: ONSET, *Ṽ, IDENTNas, MAX/n/ » *n]$_{Wd}$, INTEGRITY, UNIFORMITY, IDENT[−ct], PU
(b) [ˈsɔ̃] – [ˈsɔns] – [soˈnaɾ]: ONSET, *n]$_{Wd}$, IDENTNas, MAX/n/ » *Ṽ, INTEGRITY, UNIFORMITY, IDENT[−ct], PU
(c) [ˈsɔ] – [ˈsɔns] – [soˈnaɾ]: ONSET, *Ṽ, *n]$_{Wd}$ » IDENTNas, MAX/n/, INTEGRITY, UNIFORMITY, IDENT[−ct], PU

Then *Ṽ and *n]$_{Wd}$ exchanged places to give (23b). In the third stage (23c) *Ṽ and *n]$_{Wd}$ were both ranked above IDENTNas, MAX/n/. Finally, the presence of non-alternating *n*-final words in the language shows *n]$_{Wd}$ again demoted below faithfulness (IDENTNas, MAX/n/), with the development of lexical allomorphy in those items in which *n]$_{Wd}$ originally selected vowel-final forms, as in (22)—*Ṽ is not represented here, being undominated. It is possible, of course, that vowel nasalization occurred in pre-affix forms as well as word-finally, with or without retention of [n], but I know of no evidence that this was the case.

10.2.2 *R/zero alternation*

Similar to the phenomenon of *n*/zero alternation presented in §10.2.1 is the pattern of rhotic/zero alternation found in most Catalan dialects. However, '*r*-deletion', unlike '*n*-deletion', is not represented in modern Catalan orthography. Some examples of *r*/zero alternation are given in (24) after Bibiloni (2002: 277), in eastern continental pronunciation.

(24)

Stem	Plural /+z/	F.SG /+a/	Derived form
clar [ˈkla] 'clear'	*clars* [ˈklas]	*clara* [ˈklaɾə]	*claror* [kləˈɾo] 'light'
primer [pɾiˈme] 'first'	*primers* [pɾiˈmes]	*primera* [pɾimeˈɾə]	*primerenc* [pɾiməˈɾɛŋ] 'early'
flor [ˈflɔ] 'flower'	*flors* [ˈflɔs]		*florir* [fluˈɾi] 'to flourish'
madur [məˈðu] 'ripe'	*madurs* [məˈðus]	*madura* [məˈðuɾə]	*madurar* [məðuˈɾa] 'to ripen'

The pattern of *r*/zero alternation in (24) is characteristic of Catalan dialects other than central Valencian, where orthographic final and preconsonantal *r* is realized [ɾ]. (Northern and southern Valencian do display *r*/zero alternation.) In all the varieties where *r*/zero alternation is found there are significant numbers of lexical exceptions, which are numerous in continental Catalan but few in number on the islands, especially Majorca. The number and diversity of these exceptions in continental Catalan mean that it is quite implausible to deal with *r*/zero alternation as a motivated phonological process, though phonological factors may certainly favour or hinder '*r*-deletion', alongside morphological factors. The factors affecting the lexical incidence of '*r*-deletion'—and it is reasonable to maintain this label, as it was shown in §10.2.1 that there is no general process of consonant insertion—are these:

- grammatical category
- stress pattern: oxytone or paroxytone
- number of syllables: monosyllables vs. polysyllables
- morphological structure or derivational pattern
- quality of vowel preceding /R/
- ? word frequency

In the first place, infinitives, which contain the affix /+R/, always show deletion in word-final position, whether the infinitive is monosyllabic, oxytone, or paroxytone (25). (See §11.4.2 for infinitive /+R/ followed by clitics.) Nominalized infinitives show the same pattern, thus *dinar* [diˈna] 'lunch', *saber* [səˈβɛ] 'knowledge', with a couple of exceptions: *ésser* [ˈesə] 'be.INF' alongside *ésser* [ˈesər] 'being (N)' (see (26)); *fer* [ˈfe] 'do.INF' alongside *afer* [əˈfer] 'task, affair' and *quefer* [kəˈfer] 'task'.[7]

[7] *Afer*, at least, may be borrowed from French *affaire* rather than composed of native Catalan elements.

(25) Infinitive

r-deletion	*r*-retention
fer [ˈfe] 'do', *dir* [ˈdi] 'say'	
comprar [kumˈpɾa] 'buy', *sentir* [sənˈti] 'hear', *saber* [səˈβɛ] 'know'	
créixer [ˈkɾeʃə] 'grow', *vèncer* 'overcome', etc.	

Secondly, paroxytones other than infinitives retain /R/, with perhaps only one exception (26).

(26) Paroxytone

r-deletion	*r*-retention
càntir [ˈkanti] 'pitcher' (sole exception?), cf. *cantiret* [kənti ˈɾet] 'pitcher.DIM', *cantirer* [kənti ˈɾe] 'pitcher-maker'	*almívar* 'syrup', *dòlar* 'dollar', *cadàver* 'corpse', *càncer* 'cancer', *caràcter* 'personality', *màrtir* 'martyr', *zèfir* 'zephyr', *carnívor* 'carnivore', *fòsfor* 'phosphorus', *fèmur* 'femur', etc.

Thirdly, monosyllables generally retain /R/, though there is a significant proportion of exceptions (27). The items in bold retain /R/ even in Majorca, where *r*-deletion is more common than on the mainland, and affects most monosyllables. The lists in (27) are comprehensive, if not exhaustive.

(27) Monosyllable

r-deletion	*r*-retention
clar [ˈkla] 'clear', *dur* [ˈdu] 'hard', *flor* [flɔ] 'flower', *por* [ˈpo] 'fear', *pler* [ˈplɛ] 'pleasure' Total 5	*bar* 'bar', *bor* 'boron', *cer* 'steel', *clor* 'chlorine', *cor* 'choir', *cor* 'heart', *far* 'lighthouse', *fer* 'ferocious', *fur* 'law', *gir* 'turn', *llar* 'hearth', *llor* 'laurel', *llur* 'their', *mar* 'sea', *mer* 'mere', *mor* 'die.3SG.PR.IND', 'die.2SG.IMP', *mur* 'wall', *or* 'gold', *a la par* 'at par', ***per*** 'for', ***pur*** 'pure', *rar* 'rare', ***sor*** 'Sister', ***tir*** 'shooting', *tsar* 'tzar', *tur* '*Capra caucasica*', *ver* 'true' Total 27

Verb roots which are nominalized without an affix retain /R/ (28). A paradigm uniformity effect appears to be the motivation for this case. In all conjugation I verb forms stem-final /R/ is retained because all affixes are vowel-initial. For similar reasons, though stronger in principle because more tightly integrated into the verb paradigm, Balearic first person singular present indicative forms, with +ɩ suffix, also retain /R/ (*consider* 'I consider', *esper* 'I hope') despite the general preference for *r*-deletion in Balearic.

(28) Nominalized verb root +Ø

r-deletion	*r*-retention
reverber 'reverberation'	*albir* 'judgement' cf. *albirar* 'to judge', *atur* 'stoppage' cf. *aturar* 'to stop', *conjur* 'conspiracy' cf. *conjurar* 'to conspire', *decor* 'decoration' cf. *decorar* 'to decorate', *desesper* 'despair' cf. *deseperar* 'to despair', *enyor* 'longing' cf. *enyorar* 'to long', *perjur* 'perjurer' cf. *perjurar* 'to perjure', *respir* 'breathing' cf. *respirar* 'to breath', *retir* 'retirement' cf. *retirar* 'to retire', *sospir* 'sigh' cf. *sospirar* 'to sigh'; and from (27) *gir* 'turn' cf. *girar* 'to turn', *tir* 'shooting' cf. *tirar* 'to shoot'

A preceding high vowel strongly favours *r*-retention (29). The lexemes cited here illustrating *r*-deletion after /i/ have no alternating forms with [ɾ] present]; that is, only the orthography could actually suggest input /R/ to speakers. The three words with *r*-deletion after [u] have inflected (feminine) forms with [ɾ].

(29) Preceding high vowel

	r-deletion	*r*-retention
/i/	(*ahir* 'yesterday'. *monestir* 'monastery', *nadir* 'nadir', *poncir* 'citron')	*emir* 'emir', *papir* 'papyrus', *quefir* 'kefir', *safir* 'sapphire', *vampir* 'vampire', etc.
/u/	*dur* 'hard', *madur* 'ripe', *segur* 'sure'	*Artur* (given name), *atzur* 'azure', *cangur* 'kangaroo', *obscur* 'obscure', *futur* 'future', *hidrur* 'hydride', *immatur* 'immature', *manicur* 'manicurist', *silur* 'catfish', *tafur* 'cardsharper', etc.

The examples of /R/ after low mid vowels, with or without deletion, are very few. In (30) are those that I have identified.

(30) Preceding low mid vowel

	r-deletion	*r*-retention
/ɛ/	*deler* 'eagerness', *parer* 'opinion', *plaer* 'pleasure'[8]	*amper* ∼ *ampere* 'amp(ère)'
/ɔ/	*flor* 'flower'	*tresor* 'treasure'

[8] *Parer* and *plaer* are nominalized forms of obsolete infinitives; there are no derivatives with /R/, so synchronic /R/-less input is quite plausible. *Deler*, though, has several derivatives with /R/, such as *delerós* 'desirous', *delerar* 'be eager'.

After /a/ the examples split fairly neatly into two classes: nouns, with *r*-deletion, and adjectives, with *r*-retention (31). See, for example, *solar* (A) [suˈlar] 'solar', *solar* (N) [suˈla] 'building plot'. These examples also illustrate the adjectival suffix *-ar*, with regular *r*-retention, and the noun suffix *-ar* with regular *r*-deletion. Exceptional nouns with *r*-retention are more numerous than exceptional adjectives with *r*-deletion. Some cases of nouns in [-ˈar] are without doubt nominalized adjectives, such as *magiar* [məˈʒjar] 'Magyar', *militar* 'career soldier', *titular* 'title-holder, headline'. It is likely that *exemplar* 'copy' falls under this heading, as a nominalization of *exemplar* 'model (A)'.

(31) Preceding /a/

	r-deletion	*r*-retention
Nouns	*aixovar* 'trousseau', *bestiar* 'livestock', *campanar* 'bell-tower', *cebar* 'onion field', *centenar* 'round hundred', *collar* 'necklace', *llindar* 'threshold', *paladar* 'palate', *quintar* (dry measure), *senglar* 'boar', etc.	*atzar* 'chance', *autocar* 'coach', *basar* 'bazaar', *billar* 'billiards', *calamar* 'squid', *caviar* 'caviar', *exemplar* 'copy', *hangar* 'hangar', *jaguar* 'jaguar'
Adjectives	*seglar* 'lay', *senar* 'odd (number)'	*avar* 'miserly', *balear* 'Balearic', *escolar* 'school' (A), *familiar* 'familiar', *nuclear* 'nuclear', *preclar* 'illustrious', *similar* 'similar', *velar* 'velar', etc.

R-deletion is found more often than not after high mid vowels /e/ and /o/. Most words ending in *-er* /-eR/ display *r*-deletion, though most such words involve one of the several frequent nominal/adjectival *-er* suffixes, as in *arxiver* [ərʃiˈβe] 'archivist', *codonyer* 'quince tree', *feiner* 'working', *solter* 'unmarried'. In addition to these affixes, though, there are numerous roots in /-eR/, the majority of which also display *r*-deletion (32).

(32) Preceding /e/

r-deletion	*r*-retention
-er suffixes (*arxiver* 'archivist', *feiner* 'working', etc.)	*auster* 'austere', *enter* 'entire', *Ester* (given name), *ester*, 'ester', *iber* 'Iberian', *sever* 'severe'

carrer 'street', *darrer* 'last',
gener 'January', *menester* 'need',
mester 'craft', *paper* 'paper',
Roser (given name), *sencer* 'whole',
Sever (given name), *sincer* 'sincere',
somier 'sprung mattress', *xofer*
'driver', etc.

The majority of words in *-or* /-oR/ also show *r*-deletion; some extremely frequent suffixes are involved here too: 'agent' *-dor* as in *domador* [dumə'ðo] 'tamer', 'instrumental' *-dor* as in *mocador* 'handkerchief', and 'location' *-dor* as in *menjador* 'dining room' all show consistent *r*-deletion, as does *-or* forming abstract nouns from adjectives (*grogor* 'yellowness') or from nouns (*germanor* 'brotherhood'). The suffix *-or* forming abstract nouns from verbs is unpredictable. See the examples in (33), along with some other categories of nouns, and adjectives, ending in *-or*.

(33) Preceding /o/

		r-deletion	*r*-retention
(a)	Concrete nouns	*comodor* 'commodore', *equador* 'equator', *fluor* 'fluorine', *llacor* 'mud', *llavor* 'seed', *sector* 'sector', *senyor* 'gentleman', *tambor* 'drum', *voltor* 'vulture'	*astor* 'goshawk', *estor* 'linen curtain', *licor* 'liqueur', *motor* 'motor'
(b)	Abstract nouns in *-or* derived from verbs	*dolor* (F) 'grief', *escalfor* (F) 'warmth', *gelor* (F) 'coldness', *pudor* (F) 'stench', *sentor* (F) 'smell', *suor* (F) 'sweat', *tardor* (F) 'autumn' (cf. *tardar* 'become late'), *tremolor* (M) 'shiver'	*amor* (M/F) 'love', *error* (M/F) 'error', *sabor* (M/F) 'taste', *temor* (M/F) 'fear', *valor* (M/F) 'value'
(c)	Underived[9] abstract nouns	*ardor* (M) 'warmth', *calor* (F) 'heat', *candor* (M/F) 'candour', *esplendor* (M) 'splendour', *factor* (M) 'factor', *fetor* (F) 'stink', *olor* (F) 'smell', *remor* (F) 'murmur', *torpor* (F) 'torpor', *xardor* (F) 'heat'	*estupor* (M/F) 'stupor', *favor* (M) 'favour', *furor* (M/F) 'fury', *honor* (M) 'honour', *horror* (M) 'horror', *humor* (M) 'mood', *rubor* (M/F) 'blush', *sopor* (M/F) 'drowsiness'
(d)	Adjectives		
	Suppletive comparatives	*major* 'greater', *millor* 'better' *menor* 'smaller', *pitjor* 'worse'	
	in *-ior*		*anterior* 'previous', *exterior* 'external', *posterior* 'later', *superior* 'superior', etc.
	root		*sonor* 'sonorous, voiced'

[9] i.e. not derived from free roots in Catalan; the bases may turn up in other derivations, e.g. *fet-* in *fètid* 'stinking'.

The pattern of *r*-deletion in *-or* adjectives (33d) is predictable on morphological grounds. Not so the pattern in *-or* nouns (33a–c), though one can observe that *r*-retention is not found in those *-or* abstract nouns that are always feminine in gender.

The pattern in (25–33) no doubt reflects the arrested lexical diffusion of a sound change whose process was originally conditioned by a number of prosodic, phonetic, and grammatical factors, as well as geographical ones which are still in evidence, though they are not surveyed in detail here. For example, suffixed adjectives such as *familiar* 'familiar', *popular*, 'popular', or *superior* 'superior' are more likely to undergo *r*-deletion in northern and north-western Catalonia. As far as phonological description is concerned, one can say that, in the relevant dialects, infinitive /+R/ does not align with the right edge of a phonological word, with deletion being preferred to other potential repairs. For the rest, though, as in the case of '*n*-deletion', a considerable number of morphemes come with two allomorphs. The *r*-final allomorph is always preferred before vowel-initial suffixes (avoiding an ONSET violation). The alternative is favoured by *CODA or *COMPLEX CODA, but when allomorphy is not available, these constraints are inactive, lying below IDENTRhotic and/or MAXRhotic.

10.3 ALTERNATIONS INVOLVING STEM-FINAL LABIALS

Catalan morphophonemic alternations involving stem-final labial consonants provide an illustration of a classic problem for phonological analysis: that of overlapping allophones. The two alternating contexts in each set are onset position and coda position. In (34) I give two examples illustrating each pattern, together with a rough count of the number of *morphemes* displaying each type of alternation (obtained from Mascaró and Rafel 1990).

(34)

	Onset	Coda	Onset example	Coda example	Lexical frequency	'Underlying'
(a)	p	p	*esgaripar* 'to shriek' *esquerpa* 'surly.F'	*esgarip* 'shriek (N)' *esquerp* 'surly.M'	54	/p/
(b)	β	p	*estrebar* 'to rest on' *corba* 'curved.F'	*estrep* 'stirrup' *corb* 'curved.M'	20	/b/
(c)	β	w	*blava* 'blue.F' *plovia* 'rain. 3SG.PST.IMPF'	*blau* 'blue.M' *plou* 'rain. 3SG.PRS.IND'	44	? '/v/'
(d)	w	w	*riuada* 'flood' *creuar* 'to cross'	*riu* 'river' *creu* 'cross (N)'	22	/w/

(e) Ø	w	*plebea* 'plebeian.F'	*plebeu* 'plebeian.M'	26	?
		coïa 'cook. 3SG.PST.IMPF'	*cou* 'cook. 3SG.PRS.IND'		
(f) j	w	*caiem* 'fall. 1PL.PR.IND'	*cau* 'fall. 3SG.PRS.IND'	9	?
		seiem 'sit. 1PL.PRS.IND'	*seu* 'sit. 3SG.PRS.IND'		

Alternation pattern (34a) involves voiceless bilabial [p] in both contexts, and there is no reason to think that anything but 'underlying' /p/ is involved. Underlying /p/ occurs both after vowels and after sonorants ([m],[10] [l], [r]) and [s]. Alternation pattern (34b) is discussed above in §10.1.1. It involves a voiced bilabial non-strident fricative in onset position alternating with a bilabial stop in coda position, which is voiceless in most contexts by word-final voicing neutralization. Since this pattern contrasts with that of (34a), the underlying unit is regarded as contrastively voiced and bilabial, hence /b/ (see §10.1.1 (1) and §10.1.7). This pattern is also found both after vowels and after sonorants. In the remaining types, (34c–f), the alternation is found only after vowels. Alternation type (34c) is the first problematic one, in varieties lacking the voiced labio-dental fricative phoneme /v/ (§2.1.3.1). In varieties with [v], of course, lexemes of the (34c) type show [v] – [w] alternation, involving neutralization with pattern (34d) in coda position, but no further overlapping. In the majority pronunciation illustrated in (34), pattern (34c) involves no unique segment type, onset [β] being shared with type (34b) and coda [w] being shared with (34d) (and, for that matter, with (34e) and (34f)). One of the morphemes of type (34c) is the frequent adjective-forming suffix *-iu*, as in *actiu* 'active'. With the abstract possibilities of underlying representations in classical Generative Phonology, this (34c) alternation could be assigned to underlying '/v/' despite the absence of labio-dental or other realizations distinct from those found in the other types in (34) (Wheeler 1975/9: 200–02). Type (34d) involves non-alternating /w/; the lexical frequency count here includes only those cases with alternation evidence, leaving aside all items such as *brou* [ˈbrɔw] 'broth' or *correu* [kuˈrɛw] 'mail' which lack inflected or derived forms with a vowel-initial suffix. Some seven or eight lexical items do not distinctly belong to one of type (34c) or (34d). Either there is variation—the feminine of *hereu* 'heir' may be *hereva* (34c) or *hereua* (34d)—or the root has some derivatives that have prevocalic [β], others that have prevocalic [w], as is the case with *cau* 'hole, hollow' – *cavar* 'to dig' vs. *encauar* 'to put into a hole'. In some Majorcan varieties, where type (34c) involves [v] – [w] alternation, there is a tendency to merge type (34d) with it, replacing [w] with [v] in onsets, a case of strengthening in onsets which tends to promote morphological simplification (Bibiloni 2002: 283).

[10] /-mp/ and /-mb/ are realized [-m] in Catalonia; see §7.2.

Type (34e) involves alternation between [w] in codas and zero in onsets. Since nearly all the examples with [w] are masculine nouns or adjectives, one might suggest that [-w] represents an allomorph of the masculine gender marker, which would account for its absence in feminine or derived forms; but this explanation cannot be extended to the few verb [w] – Ø alternations, of which one example is given in (34e). Most of the nominal lexical items of type (34e) are relatively technical or specialized (*protozou* 'protozoon', *caldeu* 'Chaldaean'), in comparison with those of type (34d). Strictly, type (34e) contrasts also with non-alternating stems ending in a stressed vowel, examples of which were given in §10.2.1 (19). Interestingly, though, very few of the non-alternating stems ending in a stressed vowel, such as *comitè* 'committee', have any derived forms at all, and only plural /+z/ among inflected forms. Such words are predominantly borrowings from French and other languages. By default, then, as it were, stems that end in a stressed vowel when they stand before a vowel-initial affix have an allomorph ending in [w] which is used in the absence of such an affix. A form like *plebeu* [pləˈβɛw] 'plebeian.M' would be favoured over *[pləˈβɛ] by the Stress-to-Weight principle, or by ˈσSTEM-C] (21). When the stressed stem-final vowel is [ˈu], alternations such as *cru* 'raw.M' – *crua* 'raw.F', *nu* 'naked.M' – *nua* 'naked.F', *tabú* 'taboo' – *tabuístic* 'taboo (Adj)' would be pronounced the same whether the alternation pairs were interpreted as e.g. /ˈkɾu/ – /ˈkɾu+a/ (19), or as /ˈkɾuw/ – /ˈkɾu+a/ (34e), or as /ˈkɾuw/ – /ˈkɾuw+a/ (34d). Perhaps the sole clear exceptions to the default pattern would be *esquí* 'ski' – *esquiar* 'to ski' and *peroné* 'fibula' – *peroneal* 'fibular', where there is a stem-final stressed vowel and no consonant in derivatives.

Type (34f) is of a distinct character from the rest. It is much less frequent and occurs only in conjugation II verbs, and then in lexemes which are irregular for other reasons. In Wheeler (1975/9) these were treated as verbs with underlying root-final /j/, but alternation between [j] and [w] is true on the surface only for verbs of this sub-class. In nominals, /j/ is not involved in morphophonemic alternation with other segments or with zero (cf. *noi* [ˈnɔj] 'boy' – *noia* [ˈnɔjə] 'girl').

The conclusion to be drawn from the set of observations in this section is that only alternation types (34a), (34b), and (34d) are phonologically motivated; and the contrast between (34a) and (34b) is neutralized in codas (by coda obstruent voicing neutralization). In those varieties where the final segment is realized [v] – [w], type (34c) is also phonologically motivated, and there is again coda neutralization between types (34c) and (34d). Alternations (34e) and (34f) generally, and (34c) in [v]-less varieties, are not synchronically phonologically motivated and involve stem allomorphy. In (34e) the choice of the [-w] allomorph is favoured by SWP, or by ˈσSTEM-C] (21), though not as might be expected by ONSET: attested *plebea* etc. involve ONSET violations. The particular distribution of allomorphy in (34f) is not favoured by phonological principles. There are no phonological reasons why the two allomorphs in each case should not be reversed, giving *[kəˈwɛm] 'fall.1PL.PR.IND', *[ˈkaj] 'fall.3SG.PRS.IND'.

11

THE SYLLABIFICATION OF PRONOMINAL CLITICS

11.1 INTRODUCTION

The system of pronominal (anaphoric) clitics is possibly the most complex element of Catalan grammar.[1] There are fourteen clitic elements, most of which are polysemous to some degree. They frequently occur in combination, up to six at a time, in theory (Bonet 2002: 976). Combinations of more than three pronominal clitics are unusual, but most clauses in spontaneous speech contain one or two pronominal clitics. In addition to the complexities arising from polysemy, there are complexities due to the fact that clitic sequences are often not straightforwardly compositional. Some of the clitics are morphologically complex—for example, in the 3PL dative clitic represented here as *elzi* three morphemes can be clearly identified: /l/ '3rd person' /+z/ 'plural' and /+i/ 'dative/locative'. There is very considerable dialect variation, in the forms of the clitics, in their ordering, and in their syntactic functions. And—what concerns us specifically here—there is substantial contextual variation in their phonological realization, within individual dialects. In this chapter I deal mostly with the system found in Barcelona and the surrounding area. I concentrate on matters of phonological realization of given clitics or clitic sequences, ignoring entirely the syntactic and morphological questions relating to how the sequences are composed. A comprehensive survey of the form and functions of Catalan clitics, with attention to dialect variation, is given in Bonet (2002).

The phonology of the pronominal clitics in the Barcelona variety has been analysed in depth by Bonet & Lloret (2001; forthcoming; 2002) from an Optimality Theory perspective, and the account I offer here draws heavily on the insights of these authors, as it does also on their data. They themselves build on an OT foundation established by Jiménez & Todolí (1995). I depart from Bonet & Lloret in some respects, in an attempt either to make the analysis more adequate or comprehensive or to bring the account into line with analyses of other phenomena in this book.

[1] The chapter on pronominal clitics is the longest of 20 in Wheeler et al. (1999).

11.2 SINGLE CLITICS

The pronominal clitics of most dialects of Catalan appear as enclitics after an imperative, an infinitive or a gerund, and as proclitics before other, finite verb forms. (Clitics do not appear with participles.) In (1) are displayed the pronominal clitics of Catalan, with their Barcelona pronunciations. The clitics are grouped according to their phonological (and in some cases morphological) structure and behaviour. The label for each clitic has been used for convenience of identification; it does not reflect all the functions a clitic can have.[2]

(1) Pronominal clitics in Barcelona Catalan

	Clitic type	Label	Pronunciation	Citation form
(a)	C-clitics	1SG	[əm], [mə], [m]	em
		2SG	[ət], [tə], [t]	et
		3REFL	[əs], [sə], [s]	es
		PRTT	[ən], [nə], [n]	en
(b)	CC(i)-clitics	1PL	[ənz], [nzə], [zə], [ənzə], [nz]	ens
		2PL	[uz], [əwz], [wzə], [zə], [əwzə], [wz]	us
		3DAT.PL	[əlzi], [lzi]	elzi
(c)	3ACC.M clitics	3ACC.M.SG	[əl], [lu], [l]	el
		3ACC.M.PL	[əlz], [luz], [lzə], [əlzə], [lz]	els
(d)	V-clitics	LOC	[i], [j]	hi
		N(euter)	[u], [w]	ho
(e)	CV(z)-clitics	3DAT.SG	[li]	li
		3ACC.F.SG	[lə], [l]	la
		3ACC.F.PL	[ləz], [ləzə]	les

(a) C-clitics: Clitics with one consonant, and variable appearance of schwa, spelled *e*.
(b) CC(i)-clitics: Clitics with two adjacent consonants, and variable appearance of schwa plus variable consonant deletion.[3] The second person plural clitic,

[2] I include in (1) the realization with a final schwa adopted by all the clitics otherwise ending in a sibilant (*es*, *ens*, *us*, *les*, *els*) when the following verb starts with a sibilant (e.g. *ens salva* [ən.zə.ˈsal.βə] 'saves us' vs. *ens tira* [əns.ˈti.ɾə] 'throws (to) us', *les salva* [lə.zə.ˈsal.βə] 'saves them.F'). In this context, the appearance of the schwa avoids the contact between two sibilants, a clear OCP effect; see §11.5 below and Bonet & Lloret (2002).

[3] I follow Bonet & Lloret (forthcoming) in using *elzi*, a non-standard spelling, as the citation form for the third person dative plural clitic. (Its standard form, not used in spontaneous speech, is identical to the third person accusative masculine plural: *los* ∼ *els*). *Ens* and *us* are spelt *(-)nos* and *(-)vos* in certain enclitic positions.

spelt *us*, appears with the clitics that contain two consonants. It is argued below that the *u* is underlyingly a glide, /w/, at least in most instances.

(c) 3ACC.M clitics: Third person masculine accusative non-reflexive C- and CC-clitics, with masculine gender allomorphy (i.e. Ø ~ [u]).

(d) V-clitics: Vocalic clitics, subject to regular glide formation in appropriate contexts.

(e) CV(z)-clitics: Clitics with one consonant, followed by a vowel representing a morpheme (/+i/ DAT, /+a/ F(eminine)), and with an optional following plural morph (/+z/).

This chapter focuses on the clitics with alternative syllabification, namely the (1a) C-clitics, the (1b) CC(i) clitics, and the (1c) masculine accusative clitics. Clitics of types (1d) and (1e) have an inherent vowel, which means that epenthesis is never necessary for their syllabification. Their vowel may, however, be subject to regular glide formation or elision, as discussed in Chapter 4.

The basic pattern of the contextual variation in clitic forms of types (1a–c) is set out in (2). The forms in (2a) illustrate proclisis, enclisis of C-clitics, and enclisis of CC(i) clitics after a vowel. CC(i) enclitics after a consonant are displayed separately in (2b). Schwas which are argued to be epenthetic are underlined. Consonants subject to deletion are struck through.

(2) (a) [ə] ~ Ø alternation only

Proclisis			
Clitic type	Context	Realization	
C-clitics	__#V	C	[n]: *n'imita* 'imitates some'
	__#C	ə̲C	[ə̲n]: *en tira* 'throws some'
CC(i)-clitics	__#V	ə̲CC(i)	[ə̲nz]: *ens obre* 'opens for us' [ə̲lzi]: *elzi obre* 'opens for them'
	__#C	ə̲CC(i)	[ə̲ns]: *ens tira* 'throws (to) us'[4] [ə̲lzi] *elzi tira* 'throws to them'
Enclisis			
Clitic type	Context	Realization	
C-clitics	V#__	C	[n]: *tiri'n* 'throw.3SG[5] some!'
	C#__	Cə̲	[nə̲]: *tirem-ne* 'let's throw some'
CC(i)-clitics	V#__	CC(i)	[ns]: *tiri'ns* 'throw.3SG (to) us!' [lzi]: *tiri'lzi* 'throw.3SG to them!'
	C#__	see (2b)	

[4] Throughout this chapter I assume in all candidates the assimilation of voicing in codas.

[5] The imperative forms which are morphologically third person are used in polite/deferential address style.

(b) [ə] ~ Ø alternation and/or C ~ Ø alternation in CC(i)-clitics

Enclisis			
Context	*ens*	*elzi*	*us*
/w/#__	[ti.ˈrɛwn.zə̲] *tireu-nos* ‘throw (pl.) (to) us!’	[ti.ˈrɛwl.zi] *tireu’lzi* ‘throw (pl.) to them!’	[ti.ˈrɛw.ʷzə̲] *tireu-vos* ‘throw (pl.) yourselves!’
/m/#__	[ti.ˈrɛm.ʷzə̲] *tirem-nos* ‘let’s throw ourselves’	[ti.ˈrɛ.mə̲l.zi] *tirem-elzi* ‘let’s throw to them’	[ti.ˈrɛ.mə̲ws] *tirem-vos* ‘let’s throw you (pl.)’
/z/#__	[ˈfe.zə̲ns] *fes-nos* ‘do to us!’	[ˈfe.zə̲l.zi] *fes-elzi* ‘do to them!’	non-existent
/nt/#__	[ti.ˈranʲ.ʷzə̲] *tirant-nos* ‘throwing (to) us’	[ti.ˈran.tə̲l.zi] *tirant-elzi* ‘throwing to them’	[ti.ˈran.tə̲ws] *tirant-vos* ‘throwing (to) you (pl.)’
/R/#__	[ti.ˈraɾ.nzə̲] *tirar-nos* ‘to throw (to) us’	[ti.ˈraɾ.lzi] *tirar’lzi* ‘to throw to them’	[ti.ˈraɾ.wzə̲] *tirar-vos* ‘to throw (to) you (pl.)’

At least since Wheeler (1975/9) it has been assumed that the ‘underlying form’ of the Catalan C-clitics and the CC(i) clitics (1a–c) consists only of consonants. Appearances of schwa [ə] in clitics and clitic clusters, other than those schwas corresponding to the feminine gender marker /+a/, have been interpreted as epenthetic, and as driven by syllabification requirements. Bonet & Lloret (2001: 8–10; forthcoming) supply reasons why this approach is superior to one that attributes formal alternation to allomorphy. To take one example, suppose the partitive pronoun *en* had, say, two allomorphs /ən/ and /nə/. The selection of the former in *en tira* ‘throws some’ would then be an arbitrary matter, given that syllabic markedness constraints would prefer *[nə.ˈtiɾə], where [nə], having an onset and no coda, is a less marked syllable type than [ən]. In the epenthesis account, the site of epenthetic schwa is determined by well-founded general principles. Thus according to Bonet & Lloret (forthcoming) the input forms are taken to be those in (3). I adopt this analysis here with the modification mentioned below.

(3) Underlying forms for Barcelona pronominal clitics

	Clitic Type	Citation form	Underlying representation
(a)	C-clitics	*es, em, et, en*	/s/, /m/, /t/, /n/
(b)	CC(i)-clitics	*ens, elzi, us*	/nz/, /l+z+i/, /wz/
(c)	3rd person ACC clitics	*el, els*	/l/ ~ /l+u/, /l+z/ ~ /l+u+z/
(d)	V-clitics	*hi, ho*	/i/, /u/
(e)	CV(z)-clitics	*li, la, les*	/l+i/, /l+a/, /l+a+z/

As for the 2PL clitic *us*, the claim that it has a glide underlyingly (/wz/) is based on its phonological behaviour in non-initial position. This clitic causes the appearance of schwa when it is preceded by another consonantal clitic—*se us obren* [səw.ˈzɔ.βɾən] 'they open on you.PL' has a schwa, like *se'ns obren* [sən.ˈzɔ.βɾən] 'they open on us'—and it behaves like CC(i)-clitics in enclitic position: *tirant-vos* [ti.ˈɾan.təws] 'throwing (to) you.PL' has a schwa, just like *tirant-elzi* [ti.ˈɾan.təl.zi] 'throwing to them'. In initial proclitic position, though, *us* always surfaces with a vowel, as in *us tira* [us.ˈti.ɾə] 'throws (to) you.PL'. (Like any other initial unstressed /u/, the vowel of /uz/ may itself be subject to regular glide formation, *si us agrada* [siw.zəˈɣɾaðə] 'if it pleases you'.) If *us* had a unique underlying form with a vowel (/uz/), its syllabification behaviour should not differ from that of the vocalic clitic *ho*, /u/, which never surfaces with a schwa—cf. *s'ho creu* [su.ˈkɾɛw] 'believes it', *tirant-ho* [ti.ˈɾan.tu] 'throwing it'). On the other hand, a unique /wz/ would be expected to undergo initial schwa epenthesis in initial proclitic position—/wz#ˈtira/ *us tira* *[əws.ˈti.ɾə] 'throws (to) you.PL' parallel to /nz#ˈtira/ *ens tira* [əns.ˈti.ɾə] 'throws (to) us'. This suggests that contextually determined allomorphy /wz/ ∼ /uz/ is indeed involved, as suggested by Mascaró (1986a). This is the position I adopt here.

11.3 CLITIC CLUSTERS

The general pattern of realization of pronominal clitics set out in (2) is in broad terms repeated when more than one clitic is present in a sequence. A few examples of clusters of clitics are set out in (4) (examples from Bonet & Lloret, forthcoming)

(4)	(a) *Se li crema*	[sə̲.li.ˈkɾe.mə]	'(something) of his/hers burns'
	(b) *Ens n'imita*	[ə̲n.zə̲.ni.ˈmi.tə]	'imitates some of ours'
	(c) *Se'ns elzi crema*	[sə̲n.zə̲l.zi.ˈkɾe.mə]	'(something) of theirs burns and it affects us'
	(d) *Se us n'obren*	[sə̲w.zə̲.ˈnɔ.βɾən]	'some of yours.PL open'
	(e) Vol *quedar-se-te-me'n* tres	[kə.ˈðar.sə̲.tə̲.mə̲n]	'wants to keep three, and it affects you and me'
	(f) *Quedem-nos-les*	[kə.ˈðɛm.zə̲.ləs]	'let's keep them.F (for ourselves)'
	(g) *Quedi-se-me-la*	[ˈke.ði.sə̲.mə̲.lə]	'keep (pol.) it.F yourself for me'

Particularly to be noted is the fact that while a C-clitic standing alone before a consonant-initial verb displays edge epenthesis in Barcelona Catalan, for example *es crema* [əs.ˈkɾemə] /s#ˈkRem+a/ 'burns.INTR', such a clitic standing before another clitic beginning with a consonant has epenthesis after it, as in the (4a) example.

Though, as mentioned previously, I do not aim here to investigate the syntactic and morphological principles governing Catalan clitic forms and sequences, it is

worth raising the question whether some at least of the polysemy in clitics and clitic groups of the Barcelona variety may reflect fusion or deletion occurring for phonological reasons. There is undoubted homonymy in certain cases, as illustrated in (5a), where phonetic [li] corresponds both to /l+i/ '3SG-IO' (orthographically *li*) and to /l#i/ '3SG.M.DO' + 'locative' (orthographically *l'hi*). Similarly in (5b) [lzi] corresponds both to /l+z+i/ '3-PL-IO' and to /l+z#i/ '3.M.DO-PL' + 'locative'.

(5) (a) *li envio un paquet* 'I am sending a parcel to him/her'
l'hi envio 'I am sending it.M there'
(b) *elzi envio un paquet* 'I am sending a parcel to them'
els hi envio 'I am sending them.M.PL there'

But [li] and [lzi] are far more radically polysemous in Barcelona Catalan: [li] corresponds to any combination of third person (non-reflexive) direct object, of any gender, with third person indirect object or with locative, provided none of the objects is plural, while [lzi] corresponds to any combination of third person direct object, of any gender, with third person indirect object, provided at least one of the objects is plural.[6] Put another way, in a clitic cluster the semantic content 'third person' and 'plural' is each formally expressed only once, though each may be intended to be understood twice. These third person combinations are (mostly) distinguished in the standard written language, and also in spoken and standard written Valencian.

(6) (a) [liˈpɔzu] polysemy

Citation forms and glosses of semantic content fused in [li]	Example in standard orthographic form	Valencian version (proclitic)
el + *li* 3SG.M.DO 3SG.IO	*l'hi poso* 'I put it.M on him/her'	*li'l pose*
la + *li* 3SG.F.DO 3SG.IO	*la hi poso* 'I put it.F on him/her'	*li la pose*
ho + *li* 3SG.N.DO 3SG.IO	*li ho poso* 'I put it on him/her'	*li ho pose*
el + *hi* 3SG.M.DO LOC	*l'hi poso* 'I put it.M there'	*l'hi pose*
la + *hi* 3SG.F.DO LOC	*la hi poso* 'I put it.F there'	*la hi pose*
ho + *hi* 3SG.N.DO LOC	*l'hi poso* 'I put it.N there'	?*l'hi pose*
li + *hi* 3SG.IO LOC	*li hi poso sucre* 'I put sugar in it for him/her'	*li hi pose sucre*

(b) [əlziˈpɔzu] polysemy

Citation forms and glosses of semantic content fused in [lzi]	Example in standard orthographic form	Valencian version (proclitic)
els + *li* 3PL.M.DO 3SG.IO	*els hi poso* 'I put them.M on him/her'	*li'ls pose*

[6] In such combinations, gender only surfaces in the combination 3PL.F.DO + locative: *les hi*.

les + *li* 3PL.F.DO 3SG.IO	*les hi poso* 'I put them.F on him/her'	*li les pose*
el + *elzi* 3SG.M.DO 3PL.IO	*els el poso* 'I put it.M on them'	*li ho pose*
la + *elzi* 3SG.F.DO 3PL.IO	*els la poso* 'I put it.F on them'	*els la pose*
ho + *elzi* 3SG.N.DO 3PL.IO	*els ho poso* 'I put it.N on them'	*els ho pose*
els + *elzi* 3PL.M.DO 3PL.IO	*els els poso* 'I put them.M on them'	*els els pose*
les + *elzi* 3PL.F.DO 3PL.IO	*els les poso* 'I put them.F on them'	*els les pose*
elzi + *hi* 3PL.IO LOC	*els hi poso sucre* 'I put sugar in it for them'	*els hi pose sucre*

A few of the 'fusions' in (6) are without doubt phonological. These are the ones in which input /i/ + /i/ is realized [i], i.e. /li/ + /i/ → [li], /lzi/ + /i/ → [lzi]. This follows from the general principles governing unstressed vowel sequences, outlined in §4.4, from which the reduction of unstressed /la/ + /i/ also follows, in the Barcelona variety. The treatment of some of the remaining clitic clusters in Wheeler (1975/9: 159–60, 175) started from the premise that a phonological process reducing adjacent cases of [l] to a single [l] was involved. But from an OT perspective this approach leaves something to be desired. In the first place, the majority of the 'reduced' third person combinations would not be expected to have strictly adjacent [ll] corresponding to the input. Adjacent [ll] would only be expected for *el* (/l/) + *li* (/l+i/) and for *el* (/l/) + *elzi* (/l+z+i/). It is conceivable that [l] . . . [l] should be reduced despite not being strictly adjacent. However, nothing comparable is found in other clitic sequences, where phonologically repeated elements other than [l] surface without difficulty, as in *ens en tira* [ənzənˈtirə] 'throws some to us' with [n] . . . [n], or *us ho proposo* [uzupruˈpɔzu] 'I suggest it to you' with [u] . . . [u]. And whereas plural /z/ does not appear twice in third person clitic clusters, one cannot generally say that [z] may not appear twice, for we find, for example, *ens els envia* [ənzəlzəmˈbiə] 'sends them.M to us', or *us ens acosteu* [uzənzəkusˈtɛw] 'you approach us' (in which *acostar-se* 'approach' is a verb with an inherent reflexive pronoun). Furthermore, 'reduction' of third person clitic clusters occurs even when there is no phonological similarity between the two elements of the cluster, as is the case when neuter *ho* is involved. All in all, then, there seems no evidence that the polysemy of [li] and [lzi] in Barcelona Catalan results from phonologically driven fusion of adjacent elements. Rather, the constraint involved is morphosyntactic, to the effect that, in a clitic cluster, third person non-reflexive and plural are each expressed no more than once, and that feminine gender is not expressed in IO+DO clusters. One might be tempted to say that reference to a third person direct object is not formally expressed at all in a clitic group involving a third person indirect object, were it not for the fact that the

plurality of a direct object *is* expressed when the indirect object is singular, for example *elzi* corresponding to *els* + *li* 3PL.M.DO-3SG.IO.

11.4 SYLLABIFICATION CONSTRAINTS AFFECTING CLITICS AND CLITIC CLUSTERS

11.4.1 *Epenthesis*

I now turn to review the constraints that give rise to the effects illustrated in §§11.1–11.3. As I am claiming that epenthesis of [ə] characterizes the realization of clitics, it will be evident that some markedness constraints outrank DEP, the constraint that penalizes insertion of phonological material, or more specifically DEPV. In the simplest cases these constraints are the ones that govern syllable structure, as discussed in §3.1, i.e. *P/C (consonants are not syllable peaks—(2) in §3.1), $*M/V_{[-hi]}$ (non-high vowels are not margins—(3) in §3.1), the sonority sequencing principle SONSEQ (sonority must increase from the beginning of an onset to the nucleus of a syllable, and must decrease from the nucleus to the end of the syllable—(4) in §3.1), and MSD6, the minimum sonority distance principle (§8.2), together with the basic syllable-structure constraints ONSET, *CODA. Any vowel that is inserted to ensure that an output conforms to these constraints is necessarily the least marked vowel of the language since its content is immune to faithfulness constraints, i.e. [ə] in those (eastern) Catalan varieties where [ə] is admissible at all (likewise, [e] in western varieties). These constraints will ensure that [ə] is added to syllabify /s#ˈkRem+a/ 'burns.INTR'. However, without further constraints intervening above ONSET or *CODA, the constraints mentioned here will incorrectly favour *[sə.ˈkɾe.mə] over [əs.ˈkɾe.mə] *es crema*.

The realization *es crema* is preferred by alignment constraints requiring that clitics stand adjacent to verbs (Bonet & Lloret 2001: 18–19; 2002: 26) as in (7).

(7) (a) ALIGN(V-Cl): Align the left edge of a pronominal clitic with the right edge of V(erb)[−tense].
(b) ALIGN(Cl-V): Align the right edge of a pronominal clitic with the left edge of V(erb)[+tense].[7]

Constraints (7a) and (7b) jointly are abbreviated ALIGN(Cl/V). Many varieties of Catalan do indeed have *se crema*, and similarly with the other C-clitics of (3a). It might be that the ALIGN(Cl-V) alignment constraint active in these varieties is something like 'Align the left edge of a pronominal clitic with the left edge of a Prosodic Word including a Verb[+tense]', though such a constraint would not be sufficient for those varieties that nonetheless have 3SG.M *el* [əl] ∼ [el] or partitive

[7] Actually Bonet & Lloret's clitic alignment constraints as formulated in the papers cited here have the arguments of ALIGN reversed, but the required sense is evident from the context: *every* pronominal clitic is aligned with *some* verb. I correct the formulation here.

en [ən] ~ [en] in preconsonantal proclitic contexts, such as *el crema* 'burns it.M', *en crema* 'burns some'.

The tableaux (8) and (9) illustrate the syllabification in Barcelona Catalan of partitive *en* /n/ before and after a consonant; *en tira* 'throws some' in (8), *tirem-ne* 'let's throw some' in (9). The high-ranking syllable structure constraints are jointly abbreviated as σ-STRUC.

(8)

	/n#ˈtiɾa/	σ-STRUC	AL(Cl-V)	ONSET	*CODA	DEPV
(a)	ˈnti.ɾə	*!				
(b)	☞ə̲n.ˈti.ɾə			*	*	*
(c)	nə̲.ˈti.ɾə		*!			*

σ-STRUC, ALIGN(Cl-V) » ONSET, *CODA, DEPV

In candidate (8c), /n/ is badly aligned with its adjacent tensed verb, being separated from it by [ə]. Candidate (8b) avoids this despite its ONSET and *CODA violations.

(9)

	/tiˈrɛm#n/	σ-STRUC	AL(Cl-V)	*CODA	DEPV
(a)	ti.ˈrɛmn	*!		*	
(b)	ti.ˈrɛ.mə̲n		*!	*	*
(c)	ti.ˈrɛ.mə̲.nə̲		*!		**
(d)	☞ ti.ˈrɛm.nə̲			*	*

σ-STRUC, ALIGN(Cl-V) »*CODA, DEPV

Bonet & Lloret intend that their clitic alignment constraints should be adequate for clitic sequences as well as individual clitics. They state (2001: fn. 6; forthcoming, fn. 9) that their account uses the term 'clitic group' only descriptively. However, an account relying, as theirs does, on the conventional gradient interpretation of alignment constraints runs up against problems with examples like *se li crema* '(something) of his/hers burns' in (4a), as tableau (10) shows.

(10)

	/s#l+i#ˈkɾema/	σ-STRUC	AL(Cl-V)	ONSET	*CODA
(a)	zli.ˈkɾe.mə	*!	**		
(b)	☜ə̲z.li.ˈkɾe.mə		**	*	*
(c)	💣 sə̲.li.kɾe.mə		***!		

Gradient interpretation of ALIGN(Cl-V) prefers (10b) to (10c) because in (10b) the clitic /s/ is separated from the left edge of the finite verb form by two segments ([l] and [i]), while in (10c) it is separated by three: [ə], [l], and [i].

Two approaches to this difficulty suggest themselves. One is to appeal to constraints that align not individual clitics but clitic groups or clusters, giving this notion theoretical content. This approach is the one proposed by Jiménez & Todolí (1995: 432), who argue that two Catalan clitics are grouped in a structure [[clitic+clitic][host]] rather than *[clitic[clitic[host]]]. In this approach, then, the alignment constraints mention not *pronominal clitic* but *pronominal clitic cluster*. The other approach is to regard the evaluation of clitic alignment as a categorical yes/no affair.[8] A clitic is either precisely aligned with a host or it is not. On this basis all three candidates in (10) would receive one asterisk for ALIGN(Cl-V), because in each case /s/ fails to be aligned with the left edge of the tensed verb. Evaluation is thus passed to lower-ranked constraints, namely ONSET or *CODA in the hierarchy given here. (It will be shown later that other constraints ranked between ALIGN(Cl-V) and ONSET or *CODA are in fact the ones that crucially prefer CV syllabification of initial clitics in such cases.) I adopt Jiménez &Todolí's approach here for expository reasons, since it can be built into the constraint formulation, without requiring a non-standard interpretation of alignment constraint evaluation. Amended clitic alignment constraints follow at (11). It is understood here that the ordering of clitics within clusters is determined by syntactic principles (which vary somewhat between dialects, so that, for example, Majorcan *el te poses* 'you.SG put it.M on yourself' corresponds to Barcelona *te'l poses*).

(11) (a) ALIGN(V-Cl): Align the left edge of a pronominal clitic *cluster* with the right edge of V(erb)[−tense].
(b) ALIGN (Cl-V): Align the right edge of a pronominal clitic *cluster* with the left edge of V(erb)[+tense].

This clarification of the form of the clitic-alignment constraints removes a barrier to Bonet & Lloret's insightful interpretation of the syllabification of clitic sequences as displaying an 'emergence of the unmarked' effect (2001: §6.5; forthcoming: §6.4). To return to my earlier example, it can now be claimed that *se li crema* ((4a), (10)) is preferred to *[ə̲z.li. ˈkɾe.mə] because the latter shows gratuitous violation of the basic syllable structure markedness constraints ONSET and *CODA (while, by ALIGN(Cl-V) (11b) the whole clitic cluster is well-aligned). It is clearly also preferable in Catalan to incur multiple violations of DEPV rather than to violate syllable markedness. As Bonet & Lloret remark, a case like /kedi#s#m#l+ə/ *quedi-se-me-la* 'keep (pol.) it.F yourself for me!', with the output [ké.ði.sə̲.mə̲.lə], constitutes a case of apparently unmotivated epenthesis, since outputs with a single epenthetic vowel, like *[ˈke.ðiz.mə̲.lə] or *[ˈke.ði.sə̲m.lə], are syllabically well-formed. However, the correct realization with two epenthetic vowels avoids a *CODA violation. It also, in fact, avoids SYLCON violations (§3.1. (8)), inasmuch as [Vz.mV] and [Vm.lV] each displays a coda consonant less sonorous than a following onset, and we have seen in

[8] McCarthy (2003b) now gives good theoretical reasons against gradient evaluation in general.

Chapters 6 and 7 that SylCon is often active in Catalan. Despite this, if Onset and *Coda (or SylCon) alone were involved, ensuring maximally unmarked syllable structure, we should expect to observe all enclitic sequences being vowel-final. There are sufficient examples in (2b) and (4) to indicate that this is far from being always the case. Bonet & Lloret believe that the additional constraint favouring consonant-final enclitic clusters is an alignment constraint Final-C (12).[9]

(12) Final-C: Align(PrWd, R, Cons, R). Every prosodic word ends in a consonant (McCarthy & Prince 1994, cited by Bonet & Lloret 2001: 21).

Tableau (13) shows why Final-C outranks *Coda, in evaluating candidates for *quedi-se-me'n* 'keep.pol some of mine'. Candidates violating Align(V-Cl) are not considered.

(13)

	/ˈkedi#s#m#n/	σ-struc	Final-C	*Coda	DepV
(a)	ˈke.ðizmn	*!		*	
(b)	ˈke.ðiz.mə̲n			**!	*
(c)	☞ ˈke.ði.sə̲.mə̲n			*	**
(d)	ˈke.ði.sə̲.mə̲.nə̲		*!		***
(e)	ˈke.ði.sə̲m.nə̲		*!	*	**

In fact, though, reliance on Final-C leads Bonet and Lloret into potential difficulty with /tiɾaɾ#nz/ [tiˈɾan.zə] *tirar-nos* 'to throw us', a form for which they admit they have not yet an adequate account. For them the admitted problem is that completely faithful *[tiˈɾarns] ought to be preferred by Final-C over any epenthetic form, given that [Vrns] is a well-formed coda in Catalan, as displayed in *carns* [ˈkarns] 'meats'. I believe that Bonet & Lloret's account draws on the wrong alignment constraint to prefer *quedi-se-me'n* [ˈke.ði.sə̲.mə̲n] (13c) over *[ˈke.ði.sə̲.mə̲.nə̲] (13d). They overlook a prosodic generalization about enclitics in Barcelona Catalan. A sequence of verb plus enclitic never ends with a stressed syllable, even when this would be well-formed according to σ-Struc. Thus alongside ungrammatical *[tiˈɾarns], there are parallel structures with enclitics after various oxytone verb forms (14), all of which are realized with a final epenthetic vowel that is not forced by σ-Structure constraints.

(14) /tiˈɾaɾ#m/ *[tiˈɾarm] ≺ [tiˈɾar.mə] *tirar-me* 'to throw me' cf. *ferm* [ˈfɛɾm] 'firm'
/tiˈɾaɾ#t/ *[tiˈɾart] ≺ [tiˈɾar.tə] *tirar-te* 'to throw you.sg', cf. *part* [ˈpart] 'part'

[9] A constraint with a family resemblance to Final-C is [ˈσStem-C] mentioned in §10.2 (21).

/tiˈɾaɾ#s/ *[tiˈɾars] ≺ [tiˈɾar.sə] *tirar-se* 'to throw oneself', cf. *março* [ˈmars] 'March'

/tiˈɾaɾ#n/ *[tiˈɾarn] ≺ [tiˈɾar.nə] *tirar-ne* 'to throw some', cf. *carn* [ˈkarn] 'meat'

/tiˈɾaɾ#wz/ *[tiˈɾaws] ≺ [tiˈɾaw.zə] *tirar-vos* 'to throw you.PL', cf. *graus* [ˈgɾaws] 'degrees'

/tiˈɾɛw#nz/ *[tiˈɾɛwns] ≺ [tiˈɾɛwn.zə] *tireu-nos* 'throw.2PL.IMP us', cf. *clowns* [ˈklawns] 'clowns'

/tiˈɾant#s/ *[tiˈɾans] ≺ [tiˈɾan.sə] *tirant-se* 'throwing oneself', cf. *amants* /aˈmant+z/ [əˈmans] 'lovers'

All such forms as those in (14) illustrate the activity of the verb-clitic anti-alignment constraint given in (15).

(15) *ALIGNR(Clitic, ˈσ) (*ALR(C1-ˈσ)): The right edge of every pronominal clitic is not aligned with the right edge of the head syllable of a stressed foot.

The constraint *ALIGNR(Clitic, ˈσ), which is never violated in the Barcelona variety being investigated here (though it is in other varieties), is one of a small number of alignment constraints whose function appears to be to protect the phonological salience of the clitics and hinder their prosodic absorption into their hosts, a matter to which I return below (21).

Returning to *quedi-se-me'n* (13), I need now to show how the grammar prefers (13c) to (13d) or (13e) in the absence of Bonet & Lloret's FINAL-C constraint. There are two modifications to be proposed. The first is that DEPV, in fact, outranks *CODA: it is worse to insert a vowel than to have a coda, other things being equal. The second is that there is a constraint outranking DEPV that does penalize codas, but only internal ones, i.e. codas followed by onsets: *C]$_\sigma$C. In Chapter 7 (24) I introduced the constraint *CC]$_\sigma$C which ranks above *CC]$_\sigma$, and here *C]$_\sigma$C » *C]$_\sigma$ (= *CODA) follows the same logic. The motivation for the ranking of internal *CODA (*C]$_\sigma$C) over final *CODA (*C]$_\sigma$) is perceptual. Other things being equal, the features of a word-final coda consonant are more easily perceptible than those of an internal coda consonant, whose features are to some extent masked by those of the following onset consonant. Unlike the alignment constraint FINAL-C, then, *C]$_\sigma$C is a simple markedness constraint, which inherently outranks *CODA, being more specific. This analysis is illustrated in tableau (16).

(16)

	/ˈkedi#s#m#n/	σ-STRUC	*C]$_\sigma$C	DEPV	*CODA
(a)	ˈke.ðizmn	*!			*
(b)	ˈke.ðiz.mə̲n		*!	*	**
(c)	☞ ˈke.ði.sə̲.mə̲n			**	*
(d)	ˈke.ði.sə̲.mə̲.nə̲			***!	
(e)	ˈke.ði.sə̲m.nə̲		*!	**	*

***C]$_\sigma$C » DEPV » *CODA**

Candidates (16c) and (16e) have two DEPV violations each, and each has one syllable with a coda. They differ in that (16e)'s coda is an internal one, while (16c) has only a final coda. Candidate (16c) is preferred for this reason. The candidate with the universally least marked syllable structure (16d) is excluded for having too many DEPV violations. That (16c) wins over (16b) and (16e) also shows that a constraint requiring clitic inputs to be adjacent to each other is ranked low in Catalan. Given that σ-STRUC requirements do not allow /s/, / m/, and /n/ to stand adjacent to one another without epenthesis, (16b) and (16e) are the candidates with fewest ALIGN(Clitic-Clitic) violations. Nevertheless, the winning candidate (16c) shows the clitics worse-aligned with one another than either of those losers does. The 'no internal coda constraint' $*C]_{\sigma}C$ is, naturally, inherently outranked by 'no internal complex codas' $*CC]_{\sigma}C$ ((24) in Chapter 7). This ranking is crucial in evaluating candidates for *ens les quedem* 'we keep them.F for ourselves', as in (17), which also illustrates the activity of CONTIGUITY and MAX. Candidates violating ALIGN(Cl-V) (11b) are not considered.

(17)

	nz#l+a+z#ke'ðɛm	CONTIG	MAX	$*CC]_{\sigma}C$	$*C]_{\sigma}C$	DEPV	*CODA
(a)	zə̲.ləs.kə.'ðɛm		*!		*	*	**
(b)	ə̲z.ləs.kə.'ðɛm		*!		**	*	***
(c)	ə̲n.ləs.kə.'ðɛm		*!		**	*	***
(d)	ə̲nz.ləs.kə.'ðɛm			*!	**	*	***
(e)	nə̲zə̲.ləs.kə.'ðɛm	*!			*	**	**
(f)	☞ ə̲n.zə̲.ləs.kə.'ðɛm				*	**	***

CONTIG, MAX, $*CC]_{\sigma}$**C** » $*$**C**$]_{\sigma}$**C** » DEPV » *CODA

Candidate (17a) achieves as good syllable structure as the winner, (17f), with fewer DEPV violations, but the first consonant of the initial clitic is fatally missing (contra MAX). Candidate (17e) achieves better syllable structure than the winner (fewer codas), but only by fatally inserting schwa between segments of the clitic /nz/, incurring a CONTIGUITY violation.

I return now to examine the remaining alignment constraints that govern the realization of sequences of verb plus enclitic. Bonet & Lloret (forthcoming: §6.2 (22)) mention a constraint ALIGNR(σ) which I quote here as (18).

(18) ALIGNR(σ) (ALR(σ)): ALIGN(Lex, R, σ, R): The right edge of a lexical word (Lex) has to coincide with the right edge of a syllable.

The effect of the constraint ALIGNR(σ) is very similar to the constraint UNIFORMσ that was introduced in this book in §3.1 (15), and discussed further in §7.6. I repeat it here as (19).

(19) SYLLABIC UNIFORMITY (UNIFORMσ): The syllabic organization of elements of a phonological phrase is the same as the syllabic organization of the corresponding elements in prosodic words.

A difference between ALIGNR(σ) and UNIFORMσ is that while the latter is explicitly an output–output correspondence constraint, the level of correspondence of ALIGNR(σ) is not absolutely clear. Since a lexical word can stand on its own (by definition?) and since syllable structure constraints impose a syllabification on lexical words such that they must coincide with the end of a syllable (de facto, if not necessarily), it seems likely that ALIGNR(σ) is meant as an output–output correspondence constraint. The difference between ALIGNR(σ) and UNIFORMσ, then, comes down to the difference between 'lexical word' and 'prosodic word'. In the interpretation of prosodic hierarchy assumed in this book, a sequence involving a lexical/prosodic word and a clitic (or clitic cluster) is itself a prosodic word, so ALIGNR(σ) and UNIFORMσ are, in the end, paraphrases of each other. But we shall shortly see that such a constraint is anyway not in fact crucial.

Bonet & Lloret consider whether epenthesis in *tirin-se* /ˈtiɾin#s/ [ˈtiɾin.sə] 'throw.IMP.PL.POL yourselves!' might be accounted for by ALIGNR(σ) outranking DEPV and FINAL-C, and show that while an example of this particular structure is consistent with such a constraint hierarchy (20a), other verb-clitic clusters are not. Tableau (20) makes this clear. The other examples are (20b) *tiri'n* /ˈtiɾi#n/ [ˈtiɾin] 'throw.IMP.SG.POL some!' and (20c) *tiri'ns* /ˈtiɾi#nz/ [ˈtiɾins] 'throw.IMP.SG.POL us!'. Relevant edges are indicated with ')'; $)_L$ indicates the edge of a 'lexical word'.

(20)

		ALIGNR(σ)	FINAL-C	DEPV
(a)	ˈtiɾin#s			
	i. ˈti.ɾin$)_L$s$)_σ$	*!		
	ii. ☞ ˈti.ɾin$)_{L,σ}$.sə̲		*	*
(b)	ˈtiɾi#n			
	i. ˈti.ɾi$)_L$n$)_σ$	*!		
	ii. ☜ ˈti.ɾi$)_{L,σ}$.nə̲		*	*
(c)	ˈtiɾi#nz			
	i. ˈti.ɾi$)_L$ns$)_σ$	n s!		
	ii. ☜ ˈti.ɾi$)_L$n$)_σ$.zə̲	n	*	*

In (20c) alignment violation is calculated in a gradient way, as is conventional. The first candidate has two segments, [n] and [s], intervening between the right edge of the 'lexical word' and the right edge of a syllable. By the ranking indicated, the second candidate, better-aligned, should be preferred. Though

my analysis does not draw on FINAL-C, recourse to ALIGNR(σ) (or UNIFORMσ) is insufficient in just the same way as it is in Bonet & Lloret's account. Examples (20a) and (20c) make a crucial pair, since they contain essentially the same segmental material in the same order.

These facts lead Bonet & Lloret to conclude (2001: §6.2; forthcoming: §6.2) that what is required in the relevant alignment constraint is that the right edge of a 'lexical word' align with some syllable *constituent*: onset, nucleus, or coda. From this perspective, what is wrong with (20a.i) is that the 'lexical word' *tirin* ends in the middle of a coda. In (20b.i) and (20c.i) the right edge of *tiri* aligns with the right edge of the syllable nucleus, which is acceptable. Then, epenthesis in (20b) and (20c) is gratuitous. Bonet & Lloret call the constraint expressing this requirement ALIGN-RIGHT(sub-σ). In the formulation in (21) I modify the terminology but not the content (cf. (28) in Bonet & Lloret (forthcoming); (24) in Bonet & Lloret (2001: 22)).

(21) ALIGNR(PWd, sub-σ) (ALR(sub-σ)): ALIGN(PWd, R; M ∧ N, R): The right edge of a prosodic word (PWd) coincides with the right edge of some sub-syllabic constituent, margin (M, i.e. onset or coda) or nucleus (N).

Being more specific than ALIGNR(σ), ALIGNR(sub-σ) necessarily outranks it. It was shown in (17) that MAX and $*CC]_{\sigma}C$ outrank $*C]_{\sigma}C$. In (22) it is seen that ALIGNR(sub-σ) outranks $*C]_{\sigma}C$.

(22)

	tirin#s	ALR(sub-σ)	MAX	$*CC]_{\sigma}C$	$*C]_{\sigma}C$	DEPV	ALR(σ)
(a)	i. 'ti.ɾin)$_{V}$s	*!					
	ii. ☞ 'ti.ɾin)$_{V}$.sə̲				*	*	
(b)	'tiɾi#n						
	i. ☞ 'ti.ɾi)$_{V}$n						*
	ii. 'ti.ɾi)$_{V}$.nə̲					*!	
(c)	tiɾi#nz						
	i. ☞ 'ti.ɾi)$_{V}$ns						
	ii. 'ti.ɾi)$_{V}$n.zə̲				*!	*	

ALIGNR(sub-σ) » $*C]_{\sigma}C$, DEPV » ALIGNR(σ)

11.4.2 *Deletion of consonants and other issues in the syllabification of pronominal clitics*

Tableau (23) shows that ALIGNR(sub-σ) and *ALIGNR(Cl-ˈσ) outrank MAX; so successful candidates for verb + enclitic sequences may show deletion of consonants. The example in (23a) is *tirar-nos* /tiˈɾaR#nz/ [tiˈɾan.zə] 'to throw

us',[10] with deletion of /R/. Example (23b) is *tirant-ne* /tiˈɾant#n/ [tiˈɾan.nə], with deletion of /t/. Observe that in (23b.iii) word-final [t] is 'trapped' within an onset, thereby violating ALIGNR(sub-σ). (A [tn] onset is also excluded by the onset sonority gradient constraint MSD6, of course.) In the candidates '$)_V$' indicates the edge of the finite verb form. The ranking in (23) will be reviewed in due course ((30)).

(23)

	tiˈɾaR#nz	ALR(sub-σ)	*ALR(Cl-ˈσ)	MAX	*CC]$_\sigma$C	*C]$_\sigma$C	DEPV
(a)	i. tiˈɾar)$_V$ns	*!	*				
	ii. ☞ tiˈɾa)$_V$n.zə̲			*		*	*
	iii. tiˈɾar)$_V$n.zə̲	*!			*	*	*
(b)	tiˈɾant#n						
	i. tiˈɾan		*!				
	ii. ☞ tiˈɾan)$_V$.nə̲			*		*	
	iii. tiˈɾan.t)$_V$nə̲	*!				*	*

The other alignment constraints that play an important role in the syllabification of sequences of verb host and clitic are the two mentioned in (11) which align clitic clusters with verbs. These constraints can induce violations of the highest constraint (other than σ-STRUC) so far considered, namely ALIGNR(sub-σ). As mentioned before, *ALIGNR(Cl-ˈσ) is unviolated in Barcelona Catalan, so it also properly belongs among the constraints at the top of the hierarchy. To show the ranking ALIGN(Cl/V) » ALIGNR(sub-σ) consider the examples in (25): *tireu'lzi* /tiˈɾɛw#l+z+i/ [tiˈɾɛwl.zi] 'throw.2PL.IMP to them!'

The comparison between (25a) and (25b) is what demonstrates the ranking ALIGN(V-Cl) » ALIGNR(sub-σ). Tableau (25) also illustrates the ranking above ALIGNR(sub-σ) of two constraints not yet discussed that I spell out here in (24).

(24) (a) CONTIGUITYMorpheme (CONTIGM): Adjacent morphs within a lexical item are not separated by epenthetic material (equivalent to ALIGNMORPH in Bonet & Lloret (2001; forthcoming)). CONTIGM is a faithfulness constraint of the Contiguity family.

(b) REALIZEMorpheme (REALM): A morpheme must have some phonological exponent in the output (Bonet & Lloret 2001: 41; forthcoming: (43), after Walker 1998: 244).

[10] That [tiˈɾanzə] is the correct output, not *[tiˈɾarzə], shows that MAXNasal outranks MAXRhotic; but see (27) below.

(25)

	/tiˈɾɛw#l+z+i/	ContigM	*ALR(Cl-'σ)	RealM	AL(V-Cl)	ALR(sub-σ)	Max	*CC]σC	DepV
(a)	☞ ti.ˈɾɛw)$_V$l.zi					*		*	
(b)	ti.ˈɾɛ.w)$_V$ə̲l.zi				*!				*
(c)	ti.ˈɾɛw)$_V$.zi			*!			*		
(d)	ti.ˈɾɛw)$_V$.lə̲.zi	*!							*

ContigM, RealM, Align(V-Cl) » AlignR(sub-σ)

Example (25d), without problems of alignment and with no complex codas, nonetheless loses to (25a) inasmuch as epenthesis takes place in the loser between the morphs of /l+z+i/. Example (25c) is also well-aligned and lacks a complex internal coda, though it has a Max violation. But (25c) has not just a deleted consonant. The [l] which is deleted is the sole representative of a morpheme /+l+/. Hence this candidate is eliminated by RealizeMorpheme.

However, an apparently similar case, *tireu-nos* /tiˈɾɛw#nz/ [tiˈɾɛwn.zə] 'throw.2pl.imp us!', shows (26) that there is something wrong with the ranking AlignR(sub-σ) » Max that seemed to be established in (23). Here it gives the wrong result, favouring *[tiˈɾɛw.zə] over [tiˈɾɛwn.zə].

There is, in fact, an important difference between the deletion of consonants in the (23) examples and the absence of deletion in the required winner (26c), to which I shall return shortly. The ranking Max » AlignR(sub-σ) will prove to be the correct one.

(26)

	/tiˈɾɛ+w#nz/	ContigM	*ALR(Cl-'σ)	RealM	AL(V-Cl)	ALR(sub-σ)	Max	*CC]σC	DepV
(a)	ti.ˈɾɛw)$_V$ns		*!			*			
(b)	💣 ti.ˈɾɛw)$_V$n.zə̲					*!		*	*
(c)	☜ ti.ˈɾɛw)$_V$.zə̲						*		
(d)	ti.ˈɾɛ.w)$_V$ə̲ns				*!				*
(e)	ti.ˈɾɛn.zə			*!			*		*

In the example of (27) *tirin-nos* /ˈtiɾin#nz/ [ˈtiɾinzə] 'throw.imp.pl.pol us!' one occurrence of /n/ is elided. It is appropriate to ask whether this is strictly deletion or fusion. Given the ranking now proposed, RealM, Align(V-Cl) » Max » AlignR(sub-σ), provided Uniformity (penalizing fusion) and Max (penalizing deletion) are co-ranked, there are two tied candidates, (27e) and (27f), that are pronounced identically. In (27e) the second /n/ fuses with the following /z/; in (27f) the second /n/ is deleted. In (27) '$)_V$' indicates the edge of the finite verb form and '$(_C$' the left edge of the clitic.

(27)

	ˈtiɾi+n_1#n_2z_3	σ-STRUC	CONTIGM	*ALR(Cl-ˈσ)	REALM	AL(V-Cl)	MAX	UNIF	ALR(sub-σ)	DEPV
(a)	ˈti.ɾin$_1$)$_V$($_C$n$_2$s$_3$	*!							*	
(b)	ˈti.ɾi)$_V$($_C$n$_2$s$_3$				*!		*			
(c)	ˈti.ɾi($_C$n$_{12}$)$_V$s$_3$					*!		*		
(d)	ˈti.ɾin$_1$)$_V$($_C$s$_3$						*		*!	
(e)	☞ ˈti.ɾin$_1$)$_V$.($_C$z$_{23}$ə̲							*		*
(f)	☞ ˈti.ɾin$_1$)$_V$.($_C$z$_3$ə̲						*			*
(g)	ˈti.ɾi)$_V$($_C$n$_2$.z$_3$ə̲				*!		*		*	*
(h)	ˈti.ɾi($_C$n$_{12}$)$_V$.z$_3$ə̲					*!				*
(i)	ˈti.ɾi.n)$_V$ə̲($_C$ns					*!				*

REALM, ALIGN(V-Cl) » MAX, UNIF » ALIGNR(sub-σ) » DEPV

Candidates (27b) and (27g) delete /n_1/, in violation of REALIZEMorph, since /+n/ marks 3PL subject person/number agreement. Candidates (27c) and (27h) fuse /n_1/ and /n_2/, avoiding violation of REALIZEMorph (and of MAX), but now the verb and the clitic are ill-aligned, violating ALIGN(V-Cl). Candidate (27i) avoids fusion or deletion, but non-edge epenthesis also leads to an alignment violation. The joint winning candidates are superior to (27d), which violates ALIGNR (sub-σ). If MAXPlace is highly ranked (see below), the fusion candidate ((27e)—[n] and [z] are both denti-alveolar) will in fact be preferred over the deletion candidate (27f).

I can now return to the case of *tireu-nos* mentioned above in (26). Tableau (28) demonstrates the difference between *tireu-nos* [tiˈɾɛwn.zə̲] ‘throw.2PL us’, with epenthesis but no deletion (28a), and *tirem-nos* [tiˈɾɛm.zə̲], with epenthesis plus deletion of /n/ (28b). Though [tiˈɾɛwn.zə̲] has an unusual complex coda [wn.] which is not found word-internally, and word-finally only in the Anglicism *clown* [ˈklawn] ‘clown’, the coda is inherently well-formed according to the Sonority Sequencing principle.

(28)

		σ-STRUC	CONTIG	*ALR(Cl-ˈσ)	REALM	AL(V-Cl)	MAX	UNIF	ALR(sub-σ)	DEPV
(a)	tiˈɾɛ+w#n_1z_2									
	i. ti.ˈɾɛw)$_V$ns			*!					*	
	ii. ☞ ti.ˈɾɛw)$_V$n.z$\underline{ə}$								*	*
	iii. ti.ˈɾɛw)$_V$.$z_2$$\underline{ə}$						*!			*
	iii. ti.ˈɾɛw)$_V$.$z_{12}$$\underline{ə}$							*!		*
	iv. ti.ˈɾɛ.w)$_V$$\underline{ə}$ns					*!				*
	v. ti.ˈɾɛn.z$\underline{ə}$				*!		*			*
(b)	tiˈɾɛ+m_1#n_2z_3									
	i. ti.ˈɾɛmnz	*!		*					*	
	ii. ti.ˈɾɛn_2.z$\underline{ə}$				*!		*			*
	iii. ti.ˈɾɛms			*!					*	
	iv. ☞ ti.ˈɾɛm_1.$z_{23}$$\underline{ə}$							*		*
	v. ☞ tiˈɾɛm_1.$z_3$$\underline{ə}$						*			*
	vi. tiˈɾɛmn.z$\underline{ə}$	*!							*	*
	vii. tiˈɾɛ.m$\underline{ə}$ns					*!				*
	viii. tiˈɾɛm.n$\underline{ə}$z		*!							*

σ-STRUC, CONTIG, *ALIGNR(Cl-ˈσ), REALM, ALIGN(V-Cl) » MAX, UNIF » ALIGNR(sub-σ) » DEPV

In (28b) [mnz] cannot form a coda+onset cluster, and because ALIGN(V-Cl) ranks higher than MAX or UNIFORMITY, candidates (28b.iv) and (28b.v), with deletion of /n/ or fusion of /nz/ as /z/, are the winners.

Tableau (29) shows essentially a comparable situation to (28b), with respect to *tireu-vos* [tiˈɾɛw.zə] 'throw.2PL yourselves', but the preference for [tiˈɾɛw.zə] over *[tiˈɾɛ.wus], without epenthesis, reveals faithfulness to the underlying glide character of /w/ by IDENT[−syl] (see §11.2, under (3)).

(29)

	ti'rɛ+w$_1$#w$_2$z	σ-STRUC	CONTIGM	ID[–syl]	*ALR(Cl-'σ)	REALM	AL(V-Cl)	MAX	UNIF	ALR(sub-σ)	DEPV
(a)	ti'rɛwwz	*!			*					*	
(b)	ti'rɛ.wus			*!							
(c)	ti'rɛw$_1$s				*!			*		*	
(d)	ti'rɛw$_{12}$s				*!				*	*	
(e)	☞ ti'rɛw$_{12}$.zə̲								*		*
(f)	☞ ti'rɛw$_1$.zə̲							*			*
(g)	ti'rɛw$_2$.zə̲					*!		*			*

ID[–syl], *ALIGNR(Cl-'σ), REALM, AL(V-Cl) » **MAX**, **UNIF** » ALIGNR(sub-σ) » DEPV

I now return to the examples *tirar-nos* [ti'ɾan.zə] 'to throw us' and *tirant-ne* [ti'ɾan.nə] 'throwing some' illustrated in (23). These examples are set out again in (30) with the constraint hierarchy established in the meanwhile.

(30)

	ti'ɾa+R#nz	σ-STRUC	CONTIG	*ALR(Cl-'σ)	REALM	AL(V-Cl)	MAX	UNIF	ALR(sub-σ)	DEPV
(a)	i. ti'rar)$_V$($_C$ns			*!					*	
	ii. ti'ra($_C$n$_{12}$)$_V$.zə̲					*!		*		*
	iii. 💣 ti'ra)$_V$($_C$n$_2$.zə̲				*!		*			*
	iv. ☜ ti'rar)$_V$($_C$n.zə̲								*	*
	v. ti'ra.ɾ)$_V$ə̲($_C$ns					*!				*
(b)	ti'ran$_1$t$_2$#n$_3$									
	i. ti'ran			*!			*	*		
	ii. ti'ran$_{12}$)$_V$.($_C$n$_3$ə̲							*!		*
	iii. 💣 ti'ran$_1$)$_V$.($_C$n$_3$ə̲						*!			*
	iv. ☜ ti'rant)$_V$.($_C$nə̲									*
	v. ti'ran.t)$_V$ə̲($_C$n					*!				*

It can be seen that REALIZEMorph and/or ALIGN(V-Cl) ensures the wrong outcome in tableau (30a). In /ti'ɾaR#nz/ /R/ represents the infinitive morpheme.

Indeed, with no other constraint intervening, and given a coda that conforms to SonSeq, *tirar-nos* /tiɾ+ˈa+R#nz/ should be *[tiˈɾarn.zə], just as *tireu-nos* /tiɾ+ˈɛ+w#nz/ is [tiˈɾɛwn.zə]. The same constraint ranking also wrongly favours *[tiˈɾant.nə] over a candidate with cluster reduction, (28b.ii) or (28b.iii).

However, a sequence of verb form plus enclitic might well be expected to be subject to output–output correspondence constraints that tend to impose the pronunciation of the citation form of an infinitive, such as *tirar* [tiˈɾa], with deletion of final /R/, or a gerund such as *tirant* [tiˈɾan] with cluster reduction of [nt] to [n]. A suggested formulation of such output–output correspondence constraints is given in (31). The role of MaxC(PWd-PWd) is seen below at (38).

(31) (a) DepC(PWd-PWd): A prosodic word *x* does not have a consonant without a correspondent in a prosodic word *y* that *x* contains.
(b) MaxC(PWd-PWd): A prosodic word *x* does not have a consonant without a correspondent in a prosodic word *y* that contains *x*.

The formulation as given in (31) of DepC(PWd-PWd) and MaxC(PWd-PWd) assumes as mentioned previously that a prosodic word and its dependent clitics itself makes up a prosodic word. (That is to say, Prosodic Word is a recursive category: $[[A]_{PWdy}+C]_{PWdx}$.) In the case of a prosodic word without clitics, of course, DepC(PWd-PWd) is equivalent to DepC(I-O). In some circumstances the form of a verb that surfaces in the presence of an enclitic is indeed not that which appears in the absence of clitics; DepC(PWd-PWd) is violable. Thus in the presence of one of the vocalic enclitics *hi* and *ho*, the /+R/ infinitive morph *is* realized, as in *tirar-hi* [tiˈɾaɾi] 'to throw there', *tirar-ho* [tiˈɾaɾu] 'to throw it.N'. Alternatives, with *r*-deletion, would violate either the constraint on a stressed vowel followed by a syllabic high vowel $*^{\prime}V.V_{[+hi]}$ that is discussed in §3.2.3.1 (*[tiˈɾa.i], *[tiˈɾa.u]), or if glide formation took place, would violate *AlignR(Cl-ˈσ) ((15) above) (*[tiˈɾaj], *[tiˈɾaw]). When the enclitic is a C-clitic, *r*-deletion is also blocked, in this case, I argue, by the Stress-to-weight principle (SWP) ((20c) in §8.5). Hence, *tirar-ne* [tiˈɾar.nə] 'to throw some' is more harmonic than *[tiˈɾa.nə], despite not copying *tirar* [tiˈɾa] precisely. When none of $*^{\prime}V.V_{[+hi]}$, *AlignR(Cl-ˈσ) or SWP makes a difference, i.e. when the final syllable of an infinitive is not stressed, then /R/-deletion can apply to an infinitive in all contexts. Such an example is *reconèixer-ho* [rəkuˈnɛʃəw] 'to recognize it.N', *reconèixer-ne* [rəkuˈnɛʃən] 'to recognize some'.[11]

DepC(PWd-PWd) is also violated in sequences involving a gerund, ending in /+nt/, where word-final /t/ is regularly lost, in cluster reduction, by *Geminate-Coda; see Chapter 7 (8). Thus, despite reduction in the bare gerund form *tirant* [tiˈɾan] 'throwing', the gerund+enclitic sequence *tirant-ho* 'throwing it.N' is realized [tiˈɾan.tu], rather than *[tiˈɾa.nu]. In these contexts of verb+enclitic too, then, there is evidence of the 'emergence of the unmarked'. The effect of the

[11] The issue of what drives the epenthesis of [ə] within the infinitive *reconèixer* /Re+koˈnɛʃ+R/, in the absence of /R/ on the surface, was touched on in §8.3 (15).

markedness constraint SWP is normally concealed by the effect of DEPC(I-O), the constraint that generally bans consonant insertion within words, or DEPC(PWd-PPhr), the output–output correspondence constraint penalizing the insertion of consonants between prosodic words in a phrase, insertion which might achieve the goal of heavy stressed syllables (so *tirar pedres* 'to throw stones' is [tiˈɾa.ˈpɛ.ðɾəs] rather than *[tiˈɾar.ˈpɛ.ðɾəs]; and *tirant eines* 'throwing tools' is [tiˈɾa.ˈnɛj.nəs] rather than *[tiˈran.ˈtɛj.nəs]). The claims in this paragraph are illustrated in tableaux (32–4). Tableau (32) shows again *tirar-nos* 'to throw us' seen above in (23b), but with its candidates here evaluated by higher-ranked constraints.

(32)

tirar-nos /tiˈɾa+R#nz/ [tiˈɾa]$_{PWd}$ …	*ALR(Cl-ˈσ)	SWP	DEPC (PWd-PWd)	REALM	AL(V-Cl)
tiˈɾarns	*!		*		
☞ tiˈɾan.zə̲				*	
tiˈɾarn.zə̲			*!		
tiˈɾans	*!			*	
tiˈɾar.zə̲			*!		
tiˈɾa.ɾə̲ns		*!	*		*

***ALIGNR(Cl-ˈσ),** SWP, **DEPC(PWd-PWd) » REALM**

Tableau (33) displays the role of the SWP and also *ALIGNR (Cl-ˈσ), outranking DEPC(PWd-PWd), using *tirar-ne* [tiˈɾar.nə] 'to throw some' and *tirant-ho* [tiˈɾan.tu] 'throwing it.N'.

(33)

	tirar-ne /tiˈɾa+R#n/ [tiˈɾa]$_{PWd}$ …	*ALR(Cl-ˈσ)	SWP	DEPC (PWd-PWd)	REALM	AL(V-Cl)
(a)	tiˈɾarn	*!		*		
	tiˈɾa.nə		*!		*	
	☞ tiˈɾar.nə			*		
	tiˈɾa.ɾən		*!	*		*
	tiˈɾan	*!		*	*	*
(b)	*tirant-ho* /tiˈɾa+nt#u/ [tiˈɾan]$_{PWd}$ …					
	☞ tiˈɾan.tu			*		
	tiˈɾa.nu		*!			

***ALIGNR(Cl-ˈσ), SWP » DEPC(PWd-PWd)** » REALM

Tableau (34) illustrates the pattern in which infinitive /+R/ surfaces before a V-clitic after a stressed vowel (29a) but not after an unstressed vowel (29b). The constraints that outrank SWP are illustrated.

(34)

		DEPC(I-O)	*ˈV.V[+hi]	*ALR(Cl-ˈσ)	SWP	DEPC(PWd-PWd)	REALM	AL(V-Cl)
(a)	*tirar-ho* /tiˈɾa+R#u/ [tiˈɾa]PWd …							
	☞ tiˈɾa.ɾu				*	*		
	tiˈɾa.u		*!		*		*	
	tiˈɾaw			*!			*	
	tiˈɾaɾ.tu	*!				*		*
(b)	*reconèixer-ho* /rekoˈnɛʃ+R#u/ [rəkuˈnɛʃə]PWd…							
	rəkuˈnɛ.ʃə.ɾu					*!		
	☞ rəkuˈnɛ.ʃəw						*	

DEPC(I-O), *ˈV.V[+hi], *ALIGNR(Cl-ˈσ) » SWP » DEPC(PWd-PWd) » REALM

11.4.3 *Non-edge epenthesis in verb–clitic sequences*

The data in (2b) include some cases of verb+enclitic sequences in which epenthesis takes place *between* verb and clitic, in violation of ALIGN(V-Cl). The cases involve consonant-final verb forms, followed by certain CC(i)-clitics. Representative types are repeated here in (35).

(35) (a) *tirem-elzi* 'let's throw to them' [ti.ˈɾɛ.mə̲l.zi]
tirant-elzi 'throwing to them' [ti.ˈɾan.tə̲l.zi]
(b) *tirem-vos* 'let's throw you (pl.)' [ti.ˈɾɛ.mə̲ws]
tirant-vos 'throwing (to) you (pl.)' [ti.ˈɾan.tə̲ws]

At the end of §11.2 I pointed out that the enclitic form [əws] as in (35b) demonstrates that the underlying form of this allomorph of the 2PL clitic *us* contains /w/ specified as a glide. Otherwise syllable markedness constraints, and DEPV, would always prefer the realization [us] after a consonant. That is, it is only faithfulness to underlying [–syllabic] that prevents realization as [us] in these cases. (The Barcelona variety I described in Wheeler (1975/9) did indeed display [tiˈɾantus] for *tirant-vos*.) In other respects, good clitic alignment in examples like those in (35) could only be achieved by deleting some consonant. Deletion of /m/ would violate REALIZEMorph, as /+m/ realizes '1PL'; MAXPlace would also be violated.

Deletion of /l/ in /+l+z+i/ would violate MaxLateral (and also RealizeMorph); deletion of /w/ in /wz/ would violate MaxPlace, assuming that /w/ is crucially Labial and Velar. (If /w/ reflected only Labial place, then *tirem-vos* /tiˈɾɛm$_1$#w$_2$z/ realized as *[tiˈɾɛm$_{12}$.zə] ought to be acceptable.) Deletion of /t/ in *tirant-elzi* would not be sufficient to enable syllabification without epenthesis, while deletion of /nt/ would again violate MaxPlace and RealizeMorph. Deleting any of the /z/s would violate MaxSibilant. All these faithfulness constraints, therefore, outrank Align(V-Cl). Tableau (36) takes the example *tirin-elzi* [ˈtiɾi.nəl.zi] 'throw.pl.pol to them' to demonstrate, among other things, the ranking RealizeMorph » Align(V-Cl). Candidate (36d), though better-aligned than the winner (36b), fails to realize the 3pl inflection /+n/.

(36)

	tirin-elzi /ˈtiˈri+n#l+z+i/	σ-Struc	ContigM	MaxPlace	MaxLat	Id[−syl]	MaxSib	SWP	RealM	Al(V-Cl)
(a)	ˈtiɾinl.zi	*!								
(b)	☞ ˈtiɾi.nə̲l.zi									*
(c)	ˈtiɾin.lə̲.zi		*!							
(d)	ˈtiɾil.zi								*!	
(e)	ˈtiɾin.zi				*!				*	
(f)	ˈtiɾin.li						*!		*	

σ-Struc, ContigM, MaxPlace, **MaxLat, MaxSib** » SWP » **RealM » Al(V-Cl)**

Tableaux (37–40) illustrate candidate evaluation for the four examples of (35) respectively.

(37)

tirem-elzi /tiˈɾɛ+m#l+z+i/	σ-Struc	ContigM	MaxPlace	MaxLat	Id[−syl]	MaxSib	SWP	RealM	Al(V-Cl)
tiˈɾɛml.zi	*!								
tiˈɾɛm.lzi	*!								
☞ tiˈɾɛ.mə̲l.zi							*		*
tiˈɾɛm.ləzi		*!							
tiˈɾɛm.zi				*!				*	
tiˈɾɛm.li						*!		*	
tiˈɾɛl.zi			*!					*	

σ-Struc, ContigM, MaxPlace, MaxLat, MaxSib » SWP

(38)

	tirant-elzi /ti'ɾa+nt#l+z+i/ [ti'ɾan]$_{PWd}$ …	σ-STRUC	CONTIGM	MAXC(PWd-PWd)	MAXPlace	MAXLat	ID[–syl]	MAXSib	SWP	DEPC(PWd-PWd)	REALM	AL(V-Cl)
(a)	ti'ɾant.lzi	*!								*		
(b)	ti'ɾan.lzi	*!										
(c)	☞ ti'ɾan.tə̲l.zi									*		*
(d)	ti'ɾan.zi					*!					*	
(e)	ti'ɾan.li							*!			*	
(f)	ti'ɾan.lə̲.zi		*!									
(g)	ti'ɾal.zi			*!							*	
(h)	ti'ɾa.nə̲l.zi								*!			*

σ-STRUC, CONTIGM, MAXC(PWd-PWd), MAXLat, MAXSib » SWP » DEPC(PWd-PWd)

In (38) the competition between (38g), lacking [n] that is present in the citation form of the gerund, and the winner, (38c), which has an extra [t] (faithful to the input) that is not present in the citation form of the gerund, demonstrates the ranking MAXC(PWd-PWd) » DEPC(PWd-PWd). In neither candidate is the citation form of the infinitive respected, but it is worse to delete a consonant (or fuse it) than to insert one which happens to be faithful to the input.

(39)

tirem-vos /ti'ɾɛ+m#wz/	σ-STRUC	CONTIGM	MAXC(PWd-PWd)	MAXPlace	MAXLat	ID[–syl]	MAXSib	SWP	DEPC(PWd-PWd)	REALM	AL(V-Cl)
ti'ɾɛmws	*!										
ti'ɾɛmus						*!		*			
ti'ɾɛm.zə̲				*!							*
ti'ɾɛw.zə̲			*!	*						*	
☞ ti'ɾɛ.mə̲ws								*			*

σ-STRUC, CONTIGM, **MAXPlace,** MAXLat, **ID[–syl],** MAXSib **» SWP »** DEPC(PWd-PWd)

(40)

	tirant-vos /tiˈra+nt#wz/ [tiˈran]$_{PWd}$ …	σ-STRUC	CONTIGM	MAXC(PWd-PWd)	MAXPlace	MAXLat	ID[–syl]	MAXSib	SWP	DEPC(PWd-PWd)	REALM	AL(V-Cl)
(a)	tiˈrantws	*!								*		
(b)	tiˈran.tus						*!			*		
(c)	☞ tiˈran.tə̲ws									*		*
(d)	tiˈra.nus						*!		*			
(e)	tiˈran.zə̲				*!							
(f)	tiˈraw.zə̲			*!	*						*	
(g)	tiˈra.nə̲ws								*!			*

σ-STRUC, CONTIGM, **MAXPlace**, MAXLat, **ID[–syl],** MAXSib » **SWP** » **DEPC(PWd-PWd)**

In (40), (40c) [tiˈrantəws] ≻ (40e) * [tiˈran.zə], demonstrates the ranking MAX-Place » DEPC(PWd-PWd).

11.5 EPENTHESIS IN CLITIC CLUSTERS TO AVOID GEMINATE SIBILANTS

There are cases of schwa epenthesis in clitic+verb clusters in Barcelona Catalan not so far accounted for. Some examples are given in (41).

(41) *se suposa* /s#suˈpɔza/ [sə̲.su.ˈpɔzə] ‘3SG.REFL supposes’, i.e. ‘it is supposed’
ens sembla /nz#sɛmbla/ [ə̲n.zə̲.ˈsɛmblə] ‘seems to us’
els sent /l+z#sent/ [ə̲l.zə̲.ˈsen] ‘hears them.M’
les supera /l+a+z#suˈpeɾa/ [lə.zə̲.su.ˈpeɾə] ‘overcomes them.F’

The constraints ALIGN(Cl-V) and DEPV, together with high-ranking *GEMSib, the constraint penalizing geminate sibilants discussed in §6.5.1, lead us to expect, for the examples in (41), [ə.su.ˈpɔzə], [ən.ˈsɛmblə], and [lə.su.ˈpeɾə] respectively. (These forms are, indeed, found in some varieties for which Bonet & Lloret (2002) offer an account, alongside those of the Barcelona variety exemplified in (41)). They also account for an intermediate variety which corresponds more precisely to the orthographic forms, i.e. [sə̲.su.ˈpɔzə], alongside [ən.ˈsɛmblə] and [lə.su.ˈpeɾə].) So evidently, for the forms in (41) some other constraint outranks ALIGN(Cl-V).

Non-pronominal proclitics, of similar form to some of those shown in (41), namely the plural definite articles *els* /l+z/ (M) and *les* /l+a+z/ (F), never show epenthesis before an NP host element (42).

(42) *els sons* /l+z#ˈsɔn+z/ [ə̲l.ˈsɔns] 'the sounds' *[ə̲l.zə̲.ˈsɔns]
les senyes /l+a+z#ˈsɛɲ+a+z/ [lə.ˈsɛɲəs] 'the signs' *[lə.zə̲.ˈsɛɲəs]
les zones /l+a+z#ˈzɔn+a+z/ [lə.ˈzɔnəs] 'the zones' *[lə.zə̲.ˈzɔnəs]

Observe that *GEMSib is active, as usual, in the forms of (42). The first sibilant is not realized, even though it is the exponent of a morpheme ('plural'). In Bonet & Lloret's account (2002: 26) the different behaviour of non-pronominal clitics in (42) is plausibly attributed to a clitic alignment constraint different from ALIGN (Cl-V), namely ALIGN(Cl-∼V) whose (corrected) definition I give in (43).

(43) ALIGN(Cl-∼V): Align the right edge of a determiner clitic with the left edge of a non-verb.[12]

The logic of constraint ranking requires that ALIGN(Cl-∼V) dominates whatever constraints in turn dominate ALIGN(Cl-V) in order to allow the forms in (41) to surface. The crucial intervening constraint, as argued by Bonet & Lloret, is MAXSibilant, which has been mentioned before, and whose definition is given in (44).

(44) MAXSibilant (MAXSib): Every sibilant segment of the input has a sibilant correspondent in the output; i.e. deletion of a sibilant, or fusion where the correspondent is a non-sibilant, is prohibited.[13]

One further constraint, mentioned by Bonet & Lloret (2002: 28), is relevant both to pronominal and determiner clitics. This is the output–output correspondence constraint (45) that ensures that, when adjacent input sibilants are subject to deletion or fusion, the output matches the pronunciation of the sibilant that is morpheme-initial. This constraint is, of course, reflective of positional faithfulness. Again the formulation is my own.

(45) MAXInitialC(PWd-PPhr) (MAXInitC): The initial consonant of a prosodic word has a correspondent in a phonological phrase; i.e. word-initial consonants are not deleted.

MAXInitC is formulated here in such a way as to preserve initial consonants not only in lexical bases, as in Bonet & Lloret's formulation, but also in clitics, if initial in a prosodic word. In a phrase such as *el cas s'arregla* /l#ˈkaz#s#aˈregl+a/ 'the situation is sorted out', the /z#s/ sequence may only be realized as [s] ([əlˈka.səˈreg.glə]), never as [z] (*[əlˈka.zəˈreg.glə]).

Tableau (46) shows the interaction between the constraints relevant to the examples in (41), with *els sent* 'hears them.M' taken as representative.

[12] 'Determiner' is mentioned specifically in (43) because this constraint does not govern the alignment of pronominal clitics with other pronominal clitics: *ens n'agafa* [ə̲n.zə̲.nəˈɣafə] 'takes some from us' shows vowel epenthesis between *ens* and *n'*.

[13] Bonet & Lloret's own formulation requires only that a sibilant segment have some correspondent in the output, but this formulation fails to exclude fusion, with the output displaying a non-sibilant correspondent.

(46)

	els sent /l+z$_1$#'s$_2$ent/	*GEMSib	MAXInitC	MAXSib	REALM	AL(Cl-V)	UNIF	DEPV
(a)	əls.sen	*!						*
(b)	əl.s$_2$en			*!	*			*
(c)	əl.s$_{12}$en					*	*!	*
(d)	əl.z$_1$en		*!	*				*
(e)	əl.z$_{12}$en					*	*!	*
(f)	☞ ə̲l.zə̲.sen					*		**

***GEMSib**, MAXInitC, **MAXSib** » REALM » **AL(Cl-V)** » **UNIF** » **DEPV**

The faithful candidate (46a) is excluded by *GEMSib. Candidates (46b) and (46d) each display penalized deletion (violating MAXInitC and/or MAXSib). All of the remaining candidates have a clitic misaligned with a following finite verb, in (46c) and (46e) as a result of fusion, and in (46f) as a result of vowel epenthesis. The fact that the required winner is (46f) shows that UNIFORMITY (which penalizes fusion) outranks DEPV, a ranking assumed previously but not demonstrated till now.

In contrast, (47) shows the epenthesis candidate ruled out when a determiner rather than a pronominal clitic is involved, as in *els sons* 'the sounds' from (42).

(47)

	els sons /l+z$_1$#'s$_2$ɔn+z/	*GEMSib	MAXInitC	AL(Cl-~V)	MAXSib	REALM	UNIF	DEPV
(a)	əls.sɔns	*!						*
(b)	☞ əl.s$_2$ɔns				*	*		*
(c)	əl.s$_{12}$ɔns			*!			*	*
(d)	əl.z$_1$ɔns		*!		*			*
(e)	əl.z$_{12}$ɔns			*!			*	*
(f)	ə̲l.zə̲.sɔns			*!				**

***GEMSib**, **MAXInitC**, **AL(Cl-~V)** » **MAXSib** » **REALM** » UNIF » DEPV

In (47) all misaligned candidates, whether by fusion (47c), (47e) or by epenthesis (47f), are excluded by ALIGN(Cl-∼V). There remain two candidates with deletion, (47b) and (47d). That the former wins shows that MAXInitC is active, dominating REALIZEMorph.

The variety that has epenthesis between sibilants only in the case of the C-clitic /s/ (*se suposa* [sə.su.ˈpɔzə] alongside *els sent* [əl.ˈsen], just like *els sons* [əl.ˈsɔns]—variety B in Bonet & Lloret (2002)—has only one clitic alignment constraint ALIGN(Cl-Lex) dominating MAXSibilant; i.e. in this variety clitics with verb hosts are not distinguished from clitics with non-verb hosts. ALIGN(Cl-Lex) is itself dominated, however, by a REALIZEClitic constraint that is violated by candidates in which none of the input segments of a clitic has a correspondent in the output, thereby penalizing the realization of the /s/ clitic as [ə] or Ø.

The variety that has no epenthesis between sibilants where clitics are involved (Bonet & Lloret's variety C: *se suposa* [ə̲.su.ˈpɔzə]) is of interest inasmuch as the realization of the clitic /s/ as [ə] involves some opacity. The environment prompting epenthesis, namely unsyllabifiable CC, is not present on the surface, since *GEMSib ensures deletion or fusion between adjacent sibilants. Bonet & Lloret (2002) argue that the constraint favouring [ə̲.s_{12}u.ˈpɔzə], with opaque epenthesis in their variety C, over alternatives with pure fusion *[s_{12}u.ˈpɔzə] or deletion *[s_2u.ˈpɔzə] is a paradigm uniformity constraint requiring that preverbal clitics have the form VC(C), which is, of course, the normal shape for preconsonantal C-clitics or CC-clitics on the surface. But such a constraint would appear to prefer VC shape for prevocalic clitics also, contrary to the facts. All varieties prefer *s'amaga* 'hides.REFL' [sə.ˈmaɣə] to *[ə̲.sə.ˈmaɣə]. I leave for future investigation the appropriate refinement of the account of sibilant fusion/deletion in Bonet & Lloret's C variety.

11.6 SUMMARY OF CONSTRAINT RANKING

In Table 11.1 the example competitions demonstrate only the ranking of adjacent constraints.

TABLE 11.1 Summary of constraint ranking

Constraints	Crucial competitions
σ-STRUCTURE(§11.4.1), *GEMSib (§6.5.1), CONTIGMorph (24a), MAXLateral (37), ALIGN(Cl-∼V) (43), *ALIGNR(Cl-ˈσ) (15), MAXC(PWd-PWd) (31b), MAXInitialC(PWd-PPhr) (45), CONTIGUITY (17), DEPC(I-O) (34), *ˈV.V$_{[-hi]}$ (§3.2.3.2), ID[−syl] (§11.2)	Undominated in this chapter
»	*els sons* əl.ˈs$_2$ɔns ≻ * əls.ˈsɔns, *əl.ˈs$_{12}$ɔns (47)
MAXPlace (37), MAXSibilant (37), (44)	
»	*tirem-vos* tiˈɾɛ.məws ≻ *tiˈɾɛm.zə (39) *tirem-elzi* tiˈɾɛ.məl.zi ≻ *tiˈɾɛm.l$_{23}$i (37)
SWP (§8.5 (20c))	
»	*tirant-ho* tiˈɾan.tu ≻ *tiˈɾa.nu (33b)
DEPC(PWd-PWd) (31a)	
»	*tirar-nos* tiˈɾan$_{12}$.zə ≻ * tiˈɾarn.zə (32)
REALIZEMorph (24b)	
»	*tirin-elzi* ˈtiɾi.nəl.zi ≻ *ˈtiri.lzi (36)
ALIGN(Cl-V) (7)	
»	*tirin-nos* ˈtiɾin$_1$.z$_{23}$ə ≻ *ˈtiɾi.nəns (27)
MAX (17), UNIFORMITY (27)	
»	*tireu-nos* tiˈɾɛwn.zə ≻ *tiˈɾɛw.z$_{12}$ə (28)
ALIGNR(sub-σ) (21)	
»	*tirin-nos* ˈtiɾin$_1$.z$_{23}$ə ≻ *ˈtirin$_1$s$_{23}$ (27)
*CC]$_σ$C (17)	
»	(inherent) *ens les quedem* ən.zələskəˈðɛm ≻ *ənz.ləskəˈðɛm (17)
*C]$_σ$C (16)	
»	*quedi-se-me'n* ˈkeði.səmən ≻ *ˈkeðiz.mən (16)
DEPV (§11.4.1, (8))	
»	*tiri'n* ˈtiɾin ≻ *ˈtiɾi.nə (22b)
ALIGNR(σ) (18)	

REFERENCES

ROA = Rutgers Optimality Archive http://roa.rutgers.edu/index.php3

Alarcos Llorach, Emilio (1953). El sistema fonemático del catalán, *Archivum* 3: 135–46. In Catalan in Emilio Alarcos Llorach, *Estudis de lingüística catalana* (Barcelona: Ariel, 1983), 11–31.

Alderete, John (1999). Head dependence in stress-epenthesis interaction. In Ben Hermans & Marc van Oostendorp (eds.), *The Derivational Residue in Phonological Optimality Theory* (Amsterdam: Benjamins), 29–50 (= Faithfulness to prosodic heads. ROA 94–0000).

—— (2001). *Morphologically Governed Accent in Optimality Theory* (New York: Routledge).

Astruc, Lluïsa (2003). Sentence external elements in Catalan, *Catalan Journal of Linguistics* 2: 15–31.

Badia i Cardús, Montserrat (1993). El problema del caràcter fonològic o no dels sons africats [ts], [dz], [tʃ], [dʒ], *Llengua & Literatura* 5: 307–51.

—— (2000). *Diftongs i africats, dues qüestions polèmiques de fonologia catalana* (Barcelona: Curial/Publicacions de l'Abadia de Montserrat).

—— (2001). Vocals pròpies del sistema tònic en el sistema àton del català central, *Llengua & Literatura* 12: 127–42.

Badia i Margarit, Antoni M. (1965). Problemes de la commutació consonàntica en català, *Boletim de Filologia* 21: 213–335. Repr. in Badia (1988: 247–342).

—— (1970). Les vocals tòniques *e* i *o* en el català de Barcelona, *Estudis Romànics* 12 (1963–8): 119–72. Repr. in Badia (1988: 141–90).

—— (1988). *Sons i fonemes de la llengua catalana* (Barcelona: Universitat).

Baertsch, Karen (1998). Onset sonority distance constraints through local conjunction, *Chicago Linguistic Society Papers* 34.2: 1–15.

Beckman, Jill N. (1998). Positional Faithfulness (PhD, University of Massachusetts). ROA 234–1297.

Bibiloni, Gabriel (1998). La e àtona en el català de Mallorca. In Josep Massot i Muntaner (ed.), *Estudis de llengua i literatura en honor de Joan Veny* (Barcelona: Universitat de Barcelona/Publicacions de l'Abadia de Montserrat), ii. 533–9.

—— (2002). Elisió de *-n* i *-r*, distribució de les ròtiques i altres fenòmens. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 271–85.

Boersma, Paul (1998). *Functional Phonology* (The Hague: Holland Academic Graphics, http://fonsg3.hum.uva.nl/paul/papers/funphon.pdf).

—— (1999). On the need for a separate perception grammar (MS, University of Amsterdam). ROA 358–1099.

—— (2000). The OCP in the perception grammar (MS, University of Amsterdam). ROA-435–0601.

Bonet i Alsina, Eulàlia (1984). Aproximació a l'entonació del català (tesi de llicenciatura, Bellaterra: Universitat Autònoma de Barcelona-DDL (distributed 1993)).

BONET [I ALSINA], EULÀLIA (2002). Cliticització. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 933–89.

—— & LLORET, MARIA-ROSA (1998). *Fonologia catalana* (Barcelona: Ariel).

—— —— (2001). More on alignment as an alternative to domains: the syllabification of Catalan clitics (Universitat Autònoma de Barcelona, report de recerca GGT-01–11). ROA 592–0403.

—— —— (2002). OCP effects in Catalan cliticization, *Catalan Journal of Linguistics* 1: 19–39.

—— —— (forthcoming). More on alignment as an alternative to domains: the syllabification of Catalan clitics, *Probus* (MS, Universitat Autònoma de Barcelona/Universitat de Barcelona).

—— & MASCARÓ, JOAN (1997). On the representation of contrasting rhotics. In Fernando Martínez-Gil & Alfonso Morales-Front (eds.), *Issues in the Phonology and Morphology of the Major Iberian Languages* (Washington, DC: Georgetown University Press), 103–26.

BORRÀS, JOAN-RAMON (1992). Més encara sobre les palatals valencianes. In Antoni Ferrando (ed.), *Miscel·lània Sanchis Guarner*, 2nd edn. (València, Universitat de València), i. 321–35.

BRADLEY, TRAVIS G. (2001). A typology of rhotic duration contrast and neutralization (MS, Pennsylvania State University). ROA-436-0601.

BRUGUERA I TALLEDA, JORDI (1990). *Diccionari ortogràfic i de pronúncia* (Barcelona: Enciclopèdia Catalana).

BURZIO, LUIGI (1994). *Principles of English Stress* (Cambridge: Cambridge University Press).

—— (1999). Surface-to-surface morphology: when your representations turn into constraints. ROA 341-0999.

BYBEE, JOAN (2001). *Phonology and Language Use* (Cambridge: Cambridge University Press).

CABRÉ I MONNÉ, TERESA (1993a). Interferència, ortografia i gramàtica. In Rafael Alemany, Antoni Ferrando, & Lluís B. Meseguer (eds.), *Actes del Novè Col·loqui Internacional de Llengua i Literatura Catalanes. Alacant-Elx 9–14 de setembre de 1991* (Barcelona: Publicacions de l'Abadia de Montserrat), iii. 97–109.

—— (1993b). Estructura gramatical i lexicó: el mot mínim en català (tesi doctoral, Bellaterra, Universitat Autònoma de Barcelona, Departament de Filologia Catalana).

CABRÉ [I MONNÉ], TERESA (1998). Faithfulness to prosodic edges: dialectal variation in truncated words in Catalan, *Catalan Working Papers in Linguistics* 6: 7–22.

—— & PRIETO [I VIVES], PILAR (2001). Hiatus and rising diphthong-favoring contexts in Catalan (Universitat Autònoma de Barcelona, Grup de Gramàtica Teòrica, report de recerca GGT-01-7).

—— —— (2004). Prosodic and analogical effects in lexical glide formation in Catalan, *Probus* 16.2: 113–50 (rev. version of Cabré & Prieto 2001)).

CARDONA, OSVALD (1977). *Els grups de vocals en contacte* (Barcelona: Fundació Salvador Vives Casajuana).

CASTELLANOS I LLORENÇ, CARLES, & CASTELLANOS I LLORENÇ, RAFAEL (1979). *Diccionari català–francès francès–català* (Barcelona: Enciclopèdia Catalana).

CHARLES-LUCE, JAN (1993). The effects of semantic context on voicing neutralization, *Phonetica* 50: 28–43.

—— & DINNSEN, DANIEL A. (1987). A reanalysis of Catalan devoicing, *Journal of Phonetics* 15: 187–90.

CLEMENTS, GEORGE N. (1990). The role of the sonority cycle in core syllabification. In John Kingston & Mary E. Beckman (eds.), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech* (Cambridge: Cambridge University Press), 283–333.

COLINA, SONIA (1995). A constraint-based analysis of syllabification in Spanish, Catalan, and Galician (Ph.D., University of Illinois at Urbana-Champaign, available from UMI, Ann Arbor, Mich.)

COLOMINA I CASTANYER, JORDI (1985). El parlar de la Marina Alta, *La Rella* 4: 37–54.

—— (1996) La simplificació dels grups consonàntics finals en català. In Axel Schönberger & Tilbert Dídac Stegmann (eds.), *Actes del Desè Col·loqui Internacional de Llengua i Literatura Catalanes. Frankfurt 18–25 de setembre de 1994* (Montserrat: Publicacions de l'Abadia de Montserrat), iii. 195–224.

CROFT, WILLIAM (2003). *Typology and Universals*, 2nd edn. (Cambridge: Cambridge University Press).

CROSSWHITE, KATHERINE (2004). Vowel reduction. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.), *Phonetically Based Phonology*. (Cambridge: Cambridge University Press), 191–231.

DE LACY, PAUL (2002). The formal expression of markedness (PhD, University of Massachusetts, Amherst). ROA 542-0902.

DECESARIS, JANET A[NN] (1987). Epenthesis in Catalan. In Carol Neidle & Rafael A. Núñez Cedeño (eds.), *Studies in Romance Languages: Proceedings of the 15th Linguistic Symposium on Romance Languages* (Dordrecht: Foris), 79–91.

DINNSEN, DANIEL A., & CHARLES-LUCE, JAN (1984). Phonological neutralization, phonetic implementation and individual differences, *Journal of Phonetics* 12: 49–60.

DOLS SALAS, NICOLAU A. (1993). The predictive formalization of consonantal contacts in Majorcan Catalan (empirical and theoretical bases) (M.Phil., Department of Hispanic Studies, University of Sheffield).

—— (2000). Teoria fonològica i sil·labificació. El cas del català de Mallorca (tesi doctoral, Universitat de les Illes Balears, Palma).

—— (2001). La complexitat dels marges sil·làbics. El cas del català de Mallorca, *Randa* 47: 147–56.

—— (2002). Fenòmens en grups consonàntics. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 319–43.

DOLS [SALAS], NICOLAU [A.] & WHEELER, MAX (1996). El consonantisme final del mallorquí i el 'llicenciament d'obertures', *Caplletra* 19: 51–63.

ESTEBAS VILAPLANA, EVA (2000). The use and realisation of accentual focus in Central Catalan with a comparison to English (Ph.D., University College London).

ESTEBAS-VILAPLANA, EVA (2003). The modelling of prenuclear accents in Central Catalan declaratives, *Catalan Journal of Linguistics* 2: 97–114.

FABRA, POMPEU (1908). Qüestions d'ortografia catalana. In *Primer Congrés Internacional de la Llengua Catalana* (Barcelona: Horta), 188–221. Repr.: Barcelona: Vicens Vives, 1985.

FLEMMING, EDWARD (1995). Auditory representations in phonology (Ph.D. UCLA).

GIMENO BETÍ, LLUÍS (1994). *Estudi lingüístic dels parlars de les comarques del nord de Castelló* (Castelló: Societat Castellonenca de Cultura).

GIRALT I RADIGALES, JESÚS, et al. (1993). *Diccionari de la llengua catalana* (Barcelona: Enciclopèdia Catalana).

GREEN, ANTONY DUBACH (1997). The prosodic structure of Irish, Scots Gaelic, and Manx (Ph.D., Cornell University). ROA 196–0597.

GREENBERG, JOSEPH (1978). Some generalizations concerning initial and final consonant clusters. In Joseph Greenberg, Charles A Ferguson, & Edith A. Moravcsik (eds.), *Universals of Human Language*, ii: *Phonology* (Stanford, Calif.: Stanford University Press), 243–79 (originally published 1964/5).

HARRIS, JAMES W. (1983). *Syllable Structure and Stress in Spanish: A Nonlinear Analysis* (Cambridge, Mass.: MIT Press).

HARRIS, JAMES [W.] (1993). Integrity of prosodic constituents and the domain of syllabification in Spanish and Catalan. In Kenneth Hale & Samuel Jay Keyser (eds.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger* (Cambridge, Mass.: MIT Press), 177–93.

HAYES, BRUCE (1995). *Metrical Stress Theory: Principles and Case Studies* (Chicago: Chicago University Press).

HUALDE, JOSÉ IGNACIO (1991). *Basque Phonology* (London: Routledge).

HYDE, BRETT (2001). Metrical and prosodic structure in Optimality Theory (Ph.D., Rutgers University). ROA 476–1101.

INKELAS, SHARON, ORGUN, C. ORHAN, & ZOLL, CHERYL (1997). The implication of lexical exceptions for the nature of grammar. In Iggy Roca (ed.), *Derivations and Constraints in Phonology* (Oxford: Clarendon Press), 393–418.

IVERSON, GREGORY K. (1993a). Lexical versus postlexical rule application in Catalan. In Mushira Eid & Gregory Iverson (eds.), *Principles and Prediction: The Analysis of Natural Language. Papers in Honor of Gerald Sanders* (Amsterdam: Benjamins), 339–50.

IVERSON, GREGORY [K.] (1993b). (Post)lexical rule application. In Sharon Hargus & Ellen M. Kaisse (eds.), *Studies in Lexical Phonology* (San Diego, Calif.: Academic Press), 255–76.

JIMÉNEZ [MARTÍNEZ], JESÚS (1996a). Els aplecs *ts* i *dz* en valencià. Una anàlisi des de la Teoria de l'Optimitat, *Caplletra* 19: 81–111.

—— (1996b). Problemes de la representació del fonema /ʃ/ en català occidental. In August Bover, J. Martí Olivella, & M. A. Newman (eds.), *Actes del setè Col·loqui d'Estudis Catalans a Nord-Amèrica, Berkeley, 1993* (Barcelona: Publicacions de l'Abadia de Montserrat), 125–48.

—— (1998). Valencian vowel harmony, *Rivista di Linguistica* 10: 137–61.

—— (1999) *L'estructura sil·làbica del català* (València: Institut Interuniversitari de Filologia Valenciana/Barcelona: Publicacions de l'Abadia de Montserrat).

—— (2001). L'harmonia vocàlica en valencià (MS, Universitat de València).

—— (2002). Altres fenòmens vocàlics en el mot. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 169–94.

—— & TODOLÍ, JÚLIA (1995). La forma dels pronoms clítics catalans: condicions sil·làbiques i alineament morfològic. In Carlos Martín Vide (ed.), *Actes del XIè Congrés*

de Llenguatges Naturals i Llenguatges Formals (Barcelona: Promociones y Publicaciones Universitarias), 429–37.

JULIÀ I MUNÉ, JOAN (2002). Els sons del català. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 37–87.

JUN, JONGHO (1995). Perceptual and articulatory factors in place assimilation: an optimality theoretic approach (Ph.D., UCLA).

KAGER, RENÉ (1999). *Optimality Theory* (Cambridge: Cambridge University Press).

—— (2000). Stem stress and peak correspondence in Dutch. In Joost Dekkers, Frank van der Leeuw, & Jeroen van de Weijer (eds.), *Optimality Theory: Phonology, Syntax and Acquisition* (Oxford: Oxford University Press), 121–50.

—— (2002). Rhythmic directionality by positional licensing. ROA 514-0402.

KENSTOWICZ, MICHAEL (1994). Syllabification in Chukchee: a constraints-based analysis. ROA 30–1094.

KIRCHNER, ROBERT (1998). An effort-based approach to consonant lenition (Ph.D., UCLA). ROA 276-0898.

LADEFOGED, PETER (2000). *A Course in Phonetics* (New York: Harcourt Brace).

—— & MADDIESON, IAN (1996). *Sounds of the World's Languages* (Oxford: Blackwell).

LANDMAN, MEREDITH (2002). Morphological contiguity. In Angela C. Carpenter, Andries W. Coetzee, & Paul de Lacy (eds.), *Papers in Optimality Theory II* (University of Massachusetts Occasional Papers 26; Amherst, Graduate Linguistic Student Association, University of Massachusetts), 141–69.

LINDBLOM, BJÖRN (1986). Phonetic universals in vowel systems. In John J. Ohala & Jeri J. Jaeger (eds.), *Experimental Phonology* (Orlando, Fla.: Academic Press), 13–44.

LLACH I CARLES, SÍLVIA (1998). Fonamentació fonètica de les neutralitzacions de sonoritat en català (MS, Universitat Autònoma de Barcelona, Dept de Filologia Catalana).

LLACH [I CARLES], SÍLVIA (1999). La neutralització de la sonoritat en els grups consonàntics finals de les formes verbals sense terminació del mallorquí. In *Actes del I Congrés de Fonètica Experimental, Tarragona, 22, 23 i 24 de febrer de 1999* (Barcelona: Universitat Rovira i Virgili/Universitat de Barcelona), 241–8.

LLEÓ [I PUJOL], CONCEPCIÓ [=CONXITA] (1970). *Problems of Catalan Phonology* (Studies in Linguistics and Language Learning 8, Seattle: University of Washington).

LLORET, MARIA ROSA (1992). Les africades i la representació fonològica no-lineal: estructures de contorn, *Els Marges* 46: 47–63.

—— (1997). Consonant dissimilation in the Iberian languages. In Fernando Martínez-Gil & Alfonso Morales-Front (eds.), *Issues in the Phonology and Morphology of the Major Iberian Languages* (Washington, DC: Georgetown University Press), 127–50.

—— (1998). Sobre l'estructura morfològica dels noms en català i en castellà. In Giovanni Rufino (ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza (Università di Palermo 18–24 settembre 1995)* (Tübingen: Niemeyer), ii. 557–66.

—— (2002a). Estructura sil·làbica. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 195–249.

—— (2002b). On the function of emptiness. In Bernard Fradin et al. (eds.), *Silexicales 3. Les unités morphologiques. Forum de morphologie (3e rencontres). Actes du colloque de Villeneuve d'Ascq (19–21 septembre 2002)* (Lille: Unité Mixte de Recherche 8528 du Centre Nationale de la Recherche Scientifique), 108–17.

LLORET, MARIA ROSA (2004). The phonological role of paradigms: the case of insular Catalan. In Julie Auger, J. Clancy Clements, & Barbara Vance (eds.), *Contemporary Approaches to Romance Linguistics, Selected Papers from the 33rd Linguistic Symposium on Romance Languages (LSRL), Bloomington, Indiana, April 2003* (Amsterdam: Benjamins), 275–97.

—— & VIAPLANA, JOAQUIM (1992). Les vocals finals dels elements nominals en català. In Antoni Ferrando & Albert G. Hauf (eds.), *Miscel·lània Joan Fuster: Estudis de llengua i literatura* (Barcelona: Publicacions de l'Abadia de Montserrat), v. 419–45.

LOMBARDI, LINDA (1999). Positional faithfulness and voicing assimilation in Optimality Theory, *Natural Language & Linguistic Theory* 17: 267–302.

—— (2001). Why place and voice are different: constraint-specific alternations in Optimality Theory. In Linda Lombardi (ed.), *Segmental Phonology in Optimality Theory* (Cambridge: Cambridge University Press), 13–45.

ŁUBOWICZ, ANNA (2003). Contrast preservation in phonological mappings (Ph.D. University of Massachusetts, Amherst). ROA 554–1002.

MANASTER RAMER, Alexis (1996a). A letter from an incompletely neutral phonologist, *Journal of Phonetics* 24: 477–89.

—— (1996b). Report on Alexis' dreams—bad as well as good, *Journal of Phonetics* 24: 513–19.

MARINER BIGORRA, SEBASTIÀ (1987). Retrucs teòrics de la discussió sobre fonologia de *tx*/*x* i *tj*/*j*, *Estudis de Llengua i Literatura Catalanes* 14 (*Miscel·lània Antoni M. Badia i Margarit*, 6): 109–32.

MASCARÓ, JOAN (1976/8). Catalan phonology and the phonological cycle (Ph.D., MIT, 1976; Bloomington: Indiana University Linguistics Club, 1978). Catalan version Mascaró (1983).

—— (1983). *La fonologia catalana i el cicle fonològic* (Bellaterra: Universitat Autònoma de Barcelona).

—— (1984a) Aramon – aramònic: una regla fonològica catalana no descoberta fins ara. In *Miscel·lània Aramon i Serra. Estudis de llengua i literatura catalanes oferts a R. Aramon i Serra en el seu setantè aniversari* (*Estudis Universitaris Catalans, 23–26, 1–4 de la tercera època*) (Barcelona: Curial), iv. 123–9.

—— (1984b). Continuant spreading in Basque, Catalan, and Spanish. In Mark Aronoff & Richard T. Oehrle (eds.), *Language Sound Structure: Studies in Phonology presented to Morris Halle* (Cambridge, Mass.: MIT Press), 287–98.

—— (1985). Sobre la reducció de les transformacions d'elisió. In *Estudis Gramaticals I. Working Papers in Linguistics, Universitat Autònoma de Barcelona* (Bellaterra: Universitat Autònoma de Barcelona), 197–215.

—— (1986a). *Morfologia* (Barcelona: Enciclopèdia Catalana).

—— (1986b). Compensatory diphthongization in Majorcan Catalan. In Leo Wetzels & Engin Sezer (eds.), *Studies in Compensatory Lengthening* (Dordrecht: Foris), 133–46.

—— (1987a). Underlying voicing recoverability of finally devoiced obstruents in Catalan, *Journal of Phonetics* 15: 183–6.

—— (1987b). Syllable-final processes in Catalan. In Carol Neidle & Rafael A. Núñez Cedeño (eds.), *Studies in Romance Languages: Proceedings of the 15th Linguistic Symposium on Romance Languages* (Dordrecht: Foris), 163–80.

—— (1989). On the form of segment deletion and insertion rules, *Probus* 1: 31–61.

MASCARÓ, JOAN (1991). Iberian spirantization and continuant spreading. In Albert Branchadell et al. (eds.), *Catalan Working Papers in Linguistics 1991* (Bellaterra: Universitat Autònoma de Barcelona), 167–79.

—— & RAFEL, JOAQUIM (1990). *Diccionari català invers amb informació morfològica. Catalan Reverse Dictionary with Morphological Information* (Barcelona: Publicacions de l'Abadia de Montserrat).

MCCARTHY, JOHN J. (2002). *A Thematic Guide to Optimality Theory* (Cambridge: Cambridge University Press).

—— (2003a). Richness of the Base and the determination of underlying representations (MS, University of Massachusetts, Amherst). ROA 616–0903.

—— (2003b). OT constraints are categorical. *Phonology* 20: 75–138.

—— (ed.) (2004). *Optimality Theory in Phonology: A Reader* (Oxford: Blackwell).

—— (2005). Optimal paradigms. In Laura J. Downing, T. Allan Mall, & Renate Raffelsiefen (eds.), *Paradigms in Phonological Theory* (Oxford: Oxford University Press), 170–210. ROA 485–1201.

—— & PRINCE, ALAN [S.] (1993). Generalized alignment. In Geert Booij & Jaap van Marle (eds.), *Yearbook of Morphology 1993* (Dordrecht: Kluwer), 79–153. ROA 7-0000.

—— —— (1995). Faithfulness and reduplicative identity. In Jill Beckman, Laura Walsh Dickey, & Susan Urbanczyk (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers 18; Amherst, Mass.: Graduate Linguistic Student Association, University of Massachusetts) 249–384. ROA-60-0000.

—— —— (1999). Faithfulness and identity in Prosodic Morphology. In René Kager, Harry van der Hulst, & Wim Zonneveld (eds.), *The Prosody-Morphology Interface* (Cambridge: Cambridge University Press), 218–309.

MIER, JEANNE ZANG (1986). Estudi sociolingüístic de certs aspectes de la llengua catalana, *Treballs de Sociolingüística Catalana* 6: 33–112 (English summary: A case of lexical diffusion in a contact situation, *Studia Linguistica* 41 (1987): 72–82).

MONTREUIL, JEAN-PIERRE Y. (1987a). On nuclearity. In Terrell A. Morgan, James F. Lee, & Bill VanPatten (eds.), *Language and Language Use: Studies in Spanish Dedicated to Joseph H. Matluck* (Lanham, Md.: University Presses of America), 55–68.

—— (1987b). Nucleus assignment in French and Catalan. In Ilah Fleming (ed.), *The 13th LACUS Forum, 1986* (Lake Bluff, Ill.: Linguistics Association of Canada and the United States), 146–55.

MORALES, ALFONSO (1995). On deletion rules in Catalan. In Jon Amastae, Goodall Grant, Mario Montalbetti, & Marianne Phinney (eds.), *Contemporary Research in Romance Linguistics: Papers from the 22nd LSRL El Paso/Ciudad Juárez, February 1992* (Amsterdam: Benjamins), 37–52.

NÍ CHIOSÁIN, MÁIRE, & PADGETT, JAYE (2001). Markedness, segment realization, and locality in spreading. In Linda Lombardi (ed.), *Segmental Phonology in Optimality Theory* (Cambridge: Cambridge University Press), 118–56. ROA 188-0497.

OHALA, JOHN J. (1989). Sound change is drawn from a pool of synchronic variation. In Leiv Egil Breivik & Ernst Håkon Jahr (eds.), *Language Change: Contributions to the Study of its Causes* (Berlin: Mouton de Gruyter), 173–98.

OLIVA [LLINÀS], SALVADOR (1980). *Mètrica catalana* (Barcelona: Quaderns Crema).

—— (1992). *La mètrica i el ritme de la prosa* (Barcelona: Quaderns Crema).

PADGETT, JAYE (1995a). Feature classes. In Jill Beckman, Laura Walsh Dickey, & Susan Urbanczyk (eds.), *Papers in Optimality Theory* (University of Massachusetts Occasional Papers 18; Amherst, Graduate Linguistic Student Association, University of Massachusetts, 385–420). ROA 112-0000.

—— (1995b). Partial class behavior and nasal place assimilation. ROA 113-0000.

—— (1997) Perceptual distance of contrast: vowel height and nasality. In Rachel Walker, Motoko Katayama, & Daniel Karvonen (eds.), *Phonology at Santa Cruz 5* (UC Santa Cruz, Linguistics Research Centre), 63–78.

—— (2001). Contrast dispersion and Russian palatalization. In Elizabeth Hume & Keith Johnson (eds.), *The Role of Speech Perception in Phonology* (San Diego, Calif.: Academic Press), 187–218.

—— (2003). Systemic contrast and Catalan rhotics, (MS, UC Santa Cruz). ROA 574–0203.

PALMADA [FÉLEZ], BLANCA (1993). La construcció de les síl·labes i l'elisió vocàlica del català, *Llengua & Literatura* 5: 371–92.

—— (1994a). *La fonologia del català. Els principis generals i la variació*, (Ph.D., Bellaterra: Publicacions de la Universitat Autònoma de Barcelona).

—— (1994b). Sobre la semblança excessiva en fonologia, *Els Marges* 49: 95–106.

—— (1995). From place to continuancy. In Harry van der Hulst & Jeroen van der Weijer (eds.), *Leiden in Last: HIL Phonology Papers I* (The Hague: Academic Graphics), 299–313.

—— (1997). Continuant spreading and feature organization. In Fernando Martínez-Gil & Alfonso Morales-Front (eds.), *Issues in the Phonology and Morphology of the Major Iberian Languages* (Washington, DC: Georgetown University Press), 151–71.

—— (2002). Fenòmens assimilatoris. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 251–70.

—— & SERRA, PEP (1991). On the specification of coronals. In Albert Branchadell et al. (eds.), *Catalan Working Papers in Linguistics 1991* (Bellaterra: Universitat Autònoma de Barcelona), 181–99.

PARKER, STEPHEN G. (2002). Quantifying the sonority hierarchy (Ph.D. University of Massachusetts, Amherst; published by Graduate Linguistic Student Association).

PATER, JOE (1999). Austronesian nasal substitution and other NC effects. In René Kager, Harry van der Hulst, & Wim Zonneveld (eds.), *The Prosody–Morphology Interface* (Cambridge: Cambridge University Press), 310–43.

—— (2001). Austronesian nasal substitution revisited: what's wrong with *NC (and what's not). In Linda Lombardi (ed.), *Segmental Phonology in Optimality Theory* (Cambridge: Cambridge University Press), 159–82.

PAYÀ, MARTA (2003). Prosody and pragmatics in parenthetical insertations [*sic*] in Catalan, *Catalan Journal of Linguistics* 2: 207–27.

PEREA, MARIA PILAR (2002). Flexió verbal regular. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 583–646.

PI-MALLARACH, JOSEP (1997). L'obertura de les vocals mitjanes en una mostra de neologismes del català (tesi doctoral, Universitat Autònoma de Barcelona, Bellaterra).

PIÑEROS, CARLOS-EDUARDO (2001). Non-inherent voicing and ways to fix it (MS, University of Iowa). ROA 483–1201.

PONS, CLÀUDIA (2003). The importance of being Onset. In Man van Koppen, Joanna Sio, & Mark de Vos (eds.), *Proceedings of ConSOLE X* (Leiden: Grafisch Bedrijf).

PONS I GRIERA, LÍDIA (1992). *Iodització i apitxament al Vallès: Interpretació sociolingüística i psicolingüística dels canvis fonètics* (Barcelona: Institut d'Estudis Catalans).

PORT, ROBERT F. (1996). The discreteness of phonetic elements and formal linguistics: a response to A. Manaster Ramer, *Journal of Phonetics* 24: 491–511.

PRADILLA [CARDONA], MIQUEL ÀNGEL (1998). Entonació. In Miquel Àngel Pradilla [Cardona] (ed.), *El món dels sons* (Benicarló: Alambor), 117–49.

PRADILLA CARDONA, MIQUEL-ÀNGEL (2002). Ensordiment, espirantització i fenòmens que afecten les sibilants. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 287–318.

PRIETO I VIVES, PILAR (1996). Aproximació als contorns tonals del català central, *Caplletra* 19. 161–86.

PRIETO [I VIVES], PILAR (1997). Prosodic manifestation of syntactic structure in Catalan. In Fernando Martínez-Gil & Alfonso Morales-Front (eds.), *Issues in the Phonology and Morphology of the Major Iberian Languages* (Washington, DC: Georgetown University Press), 173–94.

PRIETO I VIVES, PILAR (1999). Patrons d'associació de l'estructura tonal en català, *Catalan Working Papers in Linguistics* 7: 207–18.

PRIETO [I] VIVES, PILAR (2002a). Entonació. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 393–462.

PRIETO [I VIVES], PILAR (2002b). *Entonació. Models, teoria, mètodes* (Barcelona: Ariel).

PRINCE, ALAN, & SMOLENSKY, PAUL (1993). Optimality Theory (Rutgers University, technical report RuCCS TR-2).

RAFART I COROMINAS, SUSANNA (1999). *Diccionari de la rima* (Barcelona: Edicions 62).

RAFEL I FONTANALS, JOAQUIM (1978). Escisió i coalescència fonemàtiques: precisions teòriques. In *Estudios ofrecidos a Emilio Alarcos Llorach* (Oviedo: Universidad), ii. 195–215.

—— (1981). Dades sobre la freqüència de les unitats fonològiques en català. In *Miscel·lània Aramon i Serra. Estudis de llengua i literatura catalanes oferts a R. Aramon i Serra en el seu setantè aniversari* (*Estudis Universitaris Catalans, 23–26, 1–4 de la tercera època*) (Barcelona: Curial), ii. 473–96.

RECASENS I VIVES, DANIEL (1978). Aproximació a les cadències tonals en català, *Anuario de Filología* (Universidad de Barcelona) 3 (1977): 509–16.

—— (1986). *Estudis de fonètica experimental del català oriental central* (Montserrat: Publicacions de l'Abadia).

RECASENS [I VIVES], DANIEL (1991a). On the production characteristics of apicoalveolar taps and trills, *Journal of Phonetics* 19: 267–80.

RECASENS I VIVES, DANIEL (1991b). *Fonètica descriptiva del català (Assaig de caracterització de la pronúncia del vocalisme i consonantisme del català al segle XX)* (Barcelona: Institut d'Estudis Catalans).

RECASENS [I VIVES], DANIEL (1993). *Fonètica i fonologia* (Barcelona: Enciclopèdia Catalana).

—— & PALLARÈS, MARIA DOLORS (2000). Underlying voicing and syllabicity in Majorcan Catalan word final stop clusters. In *Proceedings of the 5th Seminar on Speech*

Production (Models and Data) (Institut für Phonetik und Sprachliche Kommunikation, University of Munich), 125–8.

Recasens [i Vives], Pallarès, Maria Dolors (2001a). Coarticulation, assimilation and blending in Catalan consonant clusters, *Journal of Phonetics* 29: 273–301.

—— —— (2001b). *De la fonètica a la fonologia* (Barcelona: Ariel).

—— —— & Fontdevila, Jordi (1997). A model of lingual coarticulation based on articulatory constraints, *Journal of the Acoustical Society of America* 102: 544–61.

Renat i Ferrís, Guillem [=Giner i Marco, Josep] (1943). Les *e* tòniques del valencià. In Joan Coromines (ed.), *Miscel·lània Pompeu Fabra. Recull de treballs de lingüística catalana i romànica* (Buenos Aires), 349–62. Repr. in Antoni Ferrando (ed.), *Josep Giner i Marco, Obra filològica (1931–1991)* (València: Denes, 1998), 61–72.

Roca, Iggy M. (1999). Stress in the Romance languages. In Harry van der Hulst (ed.), *Word Prosodic Systems in the Languages of Europe* (Berlin: Mouton de Gruyter), 659–811.

Rosenthall, Samuel (1994) Vowel/glide alternation in a theory of constraint interaction (Ph.D. University of Massachusetts, Amherst; Hamden, Conn.: Garland, 1997). ROA 126–0496.

Rossich, Albert (2000). El nom de les lletres, *Caplletra* 27: 65–86.

Ruaix i Vinyet, Josep (1986). Límits en la neutralització de *e*, *a*, *o* àtones, *Escola Catalana* 227: 3–4. Repr. in Josep Ruaix i Vinyet, *Punts conflictius de català. Deu estudis sobre normativa lingüística* (Barcelona: Barcanova, 1989), 45–51.

Salcioli Guidi, Valeria (1988). Estudio fonético-experimental de la entonación interrogativa catalana, *Estudios de Fonética Experimental* 3: 35–70.

Sanchis Guarner, Manuel (1993). *Gramàtica valenciana*, 2nd edn., ed. Antoni Ferrando (Barcelona: Alta Fulla) (1st edn. 1950).

Saragossà, Abelard (2001). La pronúncia dels numerals 17, 18 i 19 al llarg de la història, *Estudis de Llengua i Literatura Catalanes* 43 (*Miscel·lània Giuseppe Tavani* 2): 287–320.

Serra [i López], Pep (1996a). Sobre l'enduriment prematur dels conceptes: a propòsit de les erres, *Revista de Catalunya* 107: 9–31 (expanded, with OT additions, in Serra 1996b: §4.2).

Serra i López, Josep [=Pep] (1996b). La fonologia prosòdica del català (tesi de doctorat, Departament de Filologia i Filosofia, Universitat de Girona).

Smolensky, Paul (1996). The initial state and 'Richness of the Base' in Optimality Theory. ROA 154–1196.

Steriade, Donca (1994). Complex onsets as single segments: the Mazateco pattern. In Jennifer Cole & Charles Kisseberth (eds.), *Perspectives in Phonology* (Stanford, Calif.: Center for the Study of Language and Information), 203–91.

—— (1997). Phonetics in phonology: the case of laryngeal neutralization (MS, Dept. of Linguistics, UCLA <http://web.mit.edu/linguistics/www/steriade.home.html>).

—— (1999). Alternatives to the syllabic interpretation of consonantal phonotactics. In Osamu Fujimuru, Brian Joseph, & B. Palek (eds.), *Proceedings of the 1998 Linguistics and Phonetics Conference* (Prague: Karolinum Press), 205–42.

—— (2001a). Directional asymmetries in place assimilation: a perceptual account. In Elizabeth Hume & Keith Johnson (eds.), *The Role of Speech Perception in Phonology* (San Diego Calif.: Academic Press), 219–50.

—— (2001b). The phonology of perceptibility effects: the P-map and its consequences for constraint organization (MS, MIT, <http://web.mit.edu/linguistics/www/steriade.home.html>).

TIÓ [CASACUBERTA], JAUME (1987). Die Rolle der Tonhöhe in der Emphase am Beispiel des Katalanischen. In *Proceedings XIth ICPS: the Eleventh International Congress of Phonetic Sciences, August 1–7, 1987, Tallinn, Estonia, USSR* (Tallinn: Academy of Sciences of the Estonian SSR. Institute of Languages and Literature), v. 287–90.

VENNEMANN, THEO (1988). *Preference Laws for Syllable Structure* (Berlin: Mouton de Gruyter).

VENY, JOAN, & PONS I GRIERA, LÍDIA (1998). *Atles lingüístic del domini català. Etnotextos del català oriental* (Barcelona: Institut d'Estudis Catalans).

WALKER, RACHEL (1998). Nasalization, neutral segments, and opacity effects (Ph.D., UC Santa Cruz). ROA 405–0800.

WETZELS, W. LEO, & MASCARÓ, JOAN (2001). The typology of voicing and devoicing, *Language* 77: 207–44.

WHEELER, MAX W. (1975/9). Some rules in a generative phonology of modern Catalan (D.Phil., University of Oxford, 1975), published as *Phonology of Catalan* (Publications of the Philological Society 28; Oxford: Blackwell, 1979).

—— (1977). Els fonemes catalans: alguns problemes, *Els Marges* 9: 7–22.

—— (1981). Phonological variation in contemporary Catalan. In *Miscel·lània Aramon i Serra. Estudis de llengua i literatura catalanes oferts a R. Aramon i Serra en el seu setantè aniversari* (*Estudis Universitaris Catalans, 23–26, 1–4 de la tercera època;* Barcelona: Curial), ii. 603–21.

—— (1986). Catalan sandhi phenomena. In Henning Andersen (ed.), *Sandhi Phenomena in the Languages of Europe* (Berlin: Mouton de Gruyter), 475–88.

—— (1987). L'estructura fonològica de la síl·laba i del mot en català, *Estudis de Llengua i Literatura Catalanes* 14 (*Miscel·lània Antoni M. Badia i Margarit* 6): 79–108.

—— (2002). Flexió verbal irregular i verbs defectius. In Joan Solà et al. (eds.), *Gramàtica del català contemporani*, i: *Introducció. Fonètica i fonologia. Morfologia* (Barcelona: Empúries), 647–729.

—— YATES, ALAN, & DOLS [SALAS], NICOLAU [A.] (1999). *Catalan: A Comprehensive Grammar* (London: Routledge).

INDEX

n indicates a textual note, t a table.

accents: as indicators of lexical stress 8
affixes: with complex words 282–4
affricates 11–13; and alveolo-palatal fricatives 15–22
AGREE(voice) constraint 152, 161
alacantí (southern Valencian) 149, 164
alguerès (Sardinia) 2, 6
alignment constraints 4, 85, 97, 104, 285, 287, 299, 352, 353, 378, 380; resyllabification 84–8 *see also* constraints: CONTIGUITY
allomorphy: phonologically conditioned 310–38; stressed verb stems 287n
Alta Ribagorça 1, 82n
alternations: consonants/zero 327–32; stem-final labials 338–40
alveolar approximants 25
alveolar coda inputs: in Majorcan 214–15
alveolar fricatives: with half-open vowels 48; in Majorcan 206t, 207
alveolar laterals 34
alveolar taps 24, 30, 50
alveolar trills 24, 30, 50
alveolo-palatal coda inputs: in Majorcan 215–17
alveolo-palatal fricatives 13–22; in Majorcan 215
alveolo-palatal laterals 34–5
alveolo-palatal onsets 197–201
alveolo-palatal sibilants: with half-open vowels 48
anaphoric clitics *see* pronominal clitics: syllabification of 341–69
ANCHOR constraints 157
Andorra 1, 3
apitxat (*apitjat*) (pronunciation) 23, 36: voiced and voiceless obstruents 149; voiced sibilant liaison 162n
approximants: lenition 312–14
Atles Lingüístic del Domini Català (Veny & Pons) 318–24

Balearic 2, 13; clitics in 279n; diphthongs in 102; *ieisme* in 35; lenition in 322–23; mid front vowels in 73–77; Sonority Sequence in 80; stress in 37, 38, 279; verb forms with coda clusters in 269–75 *see* also Ibizan; Majorcan; Minorcan
Balearic Islands 1, 3
Barcelona Catalan: clitic clusters in 345–6, 347; fricatives in 14–15; *ieisme* in 35; pronominal clitics in 342, 344, 348–9, 351–2; voiced geminate obstruents in 12, 13; vowels in 46
Base-Derivative correspondence constraints 69, 76–7, 93
bilingualism 3
borrowings: stress 289n, 330

Camp de Tarragona 114
Capcir valley 1
Catalonia 1, 2, *map*, 3
central Catalan 2; mid front vowels in 38–41; stressed vowels in 38
*CLASH constraints 125, 126, 127, 130–1, 133, 135, 298
clitic clusters 345–69
clitics 278–9; pronominal: syllabification of 341–69
coda clusters *see* Balearic: verb forms with coda clusters in
coda-C clusters 167t; in Majorcan 206t; morpheme-internal 168–77
codas 158, 166; *COMPLEXCODA 223, 224, 229, 238, 239; and epenthesis 257–63; *GEMINATECODA 224; *THREE-MANNERCODA 229, 245–8

colon constraints 291
ColonBinarity 291
complex words: with affixes 282–4
ComplexOnset constraint 89, 101, 106, 115, 118
compound words: stress 279–81
consonant clusters: in Majorcan 207–19; realizations 190–2; reduction 220–49; sibilants 263
consonant sandhi: inflected stems before initial consonants 244–8; inflected stems before initial vowels 240–44; uninflected stems before initial consonants 237; uninflected stems before initial vowels 235–7
consonant-zero alternations 327–38
consonants 11t; classification of 189t; deletion of 355–61; and epenthesis 17, 20–21, 27, 28; gestural blending of 188, 191 *see also* alternations: stem-final labials
constraints: Agree(Voice) 161–2; Base-Derivative 69, 76–7, 93; complex words 282; *ComplexCoda 223, 224, 228, 245; *ComplexOnset 78, 101, 106, 115; Contiguity 5, 224, 244, 252, 255, 261, 356; *DistanceFront 193–4; FootBinarity 290; and formal style 177–8, 225t, 226t; *Geminate Coda 224, 226; Iamb 290, 292, 293, 294t, 295t, 296t, 306–9t; Lazy 314, 315–16, 317; Manner 245–8; MaxMorpheme 228, 229, 230; Paradigm Uniformity 230–31, 234–5, 264; prosodic words 288–97; rhythmic 297–305; sonority 250; stress: verbs 284; syllabification 348–69; VoiceGradient 160; vowel sandhi 124–5, 144t
continuants: lenition 325–7
contrasts: lenition 316–17
coronals 36, 186, 192, 204
correspondence constraints: Base-Derivative 93; vowel sandhi 124
correspondence theory: stressed vowels 93–4
cues *see* licensing-by-cue: voicing neutralization

DAC contacts: consonants 188–90, 201–4
dentals: with half-open vowels 48
denti-alveolar contacts: 192; DistanceFront constraint 194–5
denti-alveolar non-strident obstruents: in Majorcan 209–11
Derivational Consistency (syllable structure) 95, 99n, 118
devoicing *see* word-final devoicing: non-strident obstruents
diaeresis 8
diagnostic heteroglosses 2–3
digraphs: pronunciation 7
dialects 1–2, *map*; differences in 45–52
diphthongs: falling 90–1; rising 102–3, 105–6, 109–15, 119
*DistanceDAC constraints 195, 201
*DistanceFront constraint: denti-alveolar contacts 192–5
Distinct constraint 266
dorso-palatal fricatives 22–3
dorso-palatal stops 10

eastern Catalan: fricatives in 13, 14; lenition in 318–21; vowel reduction in 55, 57–9, 61–73; vowel sandhi in 139n
effort theory 314–16
El Carxe (Murcia) 1
EndRule-R constraint 282, 298
epenthesis: and clitics 348–55, 366–9; consonants 17, 20–21, 27, 28; vowels 250–75, 287–8

faithful heterorganic consonants 192
faithfulness to prosodic heads (FPH) 89, 118, 133–4, 282
falling diphthongs 90–1, 116, 119
'final *-r* deletion' 34, 333–8
FootBinarity constraint 290
Formentera 1

fortition 18
FPH *see* faithfulness to prosodic heads
France 1
fricatives 13–23, 82; allophones 10; lenition 312–15, 325, 326

geminate consonants 36–7, 202
*GEMINATECODA constraint 224, 226
*GEMINATESibilant constraint 190, 202, 214, 366
gemination 16–17, 18, 19, 21, 26; voiced stop: and Minimum Sonority Distance 265–9
gestural blending: consonants 188, 191
glide formation 88–123
glides (semivowels) 24
grave accents 8

half-close and half-open mid vowels 45–52
HEAD(PPhr) constraint 298
HEAD(PrWd) constraint 282
HEAD-DEPENDENCE constraint 288
heterorganic dental-alveolar clusters 173–4
heterorganic nasal coda clusters 173
heterorganic place clusters 168–9; in Majorcan 210
hiatus 91–5, 96, 97, 98, 99–100, 130; in Balearic 271–3
high vocoids: preceding other vocoids 99–108; sequences of 120–3
high vowel hiatus 91–6
homorganic nasal clusters 168

IAMB constraint 290, 292, 293t, 294t, 295t, 296t, 306–9t
Ibiza 1
Ibizan: alveolar taps in 25; 'homorganic stop deletion' in 220; lenition in 322–3; vowel reduction in 54 *see also* Balearic
IDENT constraints (consonants) 161, 178, 182, 192–7, 199, 202–4, 205, 223–4; in Majorcan 208, 211, 212, 215–19, 227
IDENT constraints (prominence) 298
IDENT constraints (vowels) 57, 59, 64, 67, 73
ieisme 35
INFLECTIONAL CONSISTENCY (syllable structure) 91, 118, 121
informal speech *see* speech: less formal
INITIAL CONSISTENCY: syllables 97, 118
*INITIALOVERLAP: vowel sandhi 132
INTEGRITY: in Majorcan 216–19
inter-word coda-C clusters 180–6
Italy 1

La Marina Alta 13n
labial–alveolo-palatal clusters 174
labial coda inputs: in Majorcan 211–15
labials: stem-final *see* alternations: stem-final labials
labio-dental–C clusters: in Majorcan 210
labio-velar clusters 107
labio-velar obstruents 100–102
labiodental fricatives 13
*LAPSE: rhythmic constraints 298–9, 301t, 302t, 303t, 304t, 305t; syllables 97, 98, 99, 110, 114, 115, 116–18, 119, 121; vowel sandhi 128–9, 136–7, 139, 142–3
laterals 34–37
Latin: legacy in verbal inflection 286
LAZY constraint 314, 315–16, 317
LAZYSIBILANTS 163–4
lenition 310–26
letters: pronunciation 7
licensing-by-cue: voicing neutralization 151, 152–56, 157, 158
licensing-by-prosody: voicing neutralization 151, 157

Majorca 1
Majorcan: alveolar taps in 25; cluster reduction in 221, 226–8, 232–41, 243–49; clusters in 13t, 207–219, 220; consonantal contact in 207–19; dorso-palatal fricatives in 22–3; dorso-palatal stops in 10; fricatives in 82; lenition in 321–3; obstruent voice contrasts in 150–1; verb forms with coda clusters in 269; vowel

reduction in 53, 55, 60–1, 73–6, 77t, 78t *see also* Balearic
Marked-Cluster constraints 171, 172
Mascaró's law 41–5
MaxMorpheme constraints 228, 229
Metrical Consistency: unstressed vowels 96, 98, 121, 128, 133, 134–5, 138
mid front vowels: distribution of 38–41
Mid Vowel Consistency 95, 118, 121
Minimum Sonority Distance 18, 20, 258, 262, 265–9, 313
minor place assimilation 207
Minorca 1
Minorcan: alveolar trills in 25; consonant clusters in 207; fricatives in 81; *ieisme* in 35n; lenition in 321, 323; vowel reduction in 54 *see also* Balearic
modelling variation 3–4
monosyllabic words: *r*-zero alternation in 333 *see also* clitics
moraicity: high vocoids 105
morpheme-internal coda-C clusters 168–79
morphemes: in Majorcan 207; stress 280
multi-vocoid sequences: vowel sandhi 140–43
Murcia 1
mutual assimilation: consonants *see* gestural blending: consonants

n-zero alternation 327–32
nasal and lateral assimilation: less formal style 179–80
nasalization 49
nasals 24; in Majorcan 210, 211; Sonority Sequence 79
native speakers 3
neutralization: word-final voicing 149–62
non-strident obstruents: word-final devoicing of 7–8
north Catalan 2; stress in 38, 279–80n; unstressed syllables in 45
north-western Catalan 2, 276n

obstruents 23; labio-velar 100–2; in Majorcan 209–11; non-strident 7–8; voiced 145–9, 158; voiced geminate 12–13; and voicing neutralization 159
*OCP constraints 63, 69, 126, 136
Occitan 1
Onset constraints 78–87, 112, 113, 114, 115, 116–19, 120–21; alveolo-palatal 197–201
Onset + Nucleus sequence 124
Optimal Paradigms 271–4
Optimality Theory 3–4, 5, 21–2; and voicing neutralization 157–62
orthography 6–8

P-map: voicing neutralization 164–5
palatal/velar non-strident obstruent–C clusters: in Majorcan 210
Paradigm Uniformity 76–7, 229, 231, 234, 264, 274
ParseFt constraint 291
phonological clitics 277–8
plosives 10–11
positional faithfulness 105
positional markedness 108
pretonic diphthongs 114
progressive assimilation: consonants 192
pronominal clitics: syllabification of 341–69
prosody *see* licensing by prosody: voicing neutralization
Pyrénées-Orientales 1

r-zero alternation 333–8
RealizeMorpheme constraint 356
regressive assimilation: consonants 188, 190
resyllabification 85–9, 240 *see also* syllables
rhotics 24–34, 39, 40, 190
rhythmic constraints 297–305
Ribagorça *see* Alta Ribagorça
Richness of the Base theory 93n, 250, 263
rising diphthongs 102–3, 106–7, 109–15, 118

rising sonority: vocoid sequences 103–5, 109–19
Romance languages 1
rossellonès 2 *see also* north Catalan

sandhi *see* consonant sandhi, vowel sandhi
Sardinia *see alguerès* (Sardinia)
schwas: single clitics 342n, 343, 344–5, 366
semivowels: glides 24; with half-open vowels 48
sibilants: word-final 162–4; and epenthesis 263–4
sonority: constraints 250; sonority hierarchy 253–5; sonority scale 80, 82
SONORITYSEQUENCE 18–19, 79, 83, 84, 89, 228; in Majorcan 228
southern Catalan: alveolar taps in 25; fricatives in 22
Spain 1, 3 *see also* Alta Ribagorça; Balearic Islands; Barcelona; Murcia; Valencia
Spanish: influence of 35, 39n, 52, 162n
spelling alternations 7
spirantization *see* lenition
stem-final labials: alternations 338–40
stems: cluster reduction in 221–49; in compounds 281
stress 276–309
stressed vowels: contrasts 37–52; high vowel hiatus 91–6; vowel sandhi 125–32
STRESS-TO-WEIGHT PRINCIPLE 266–7, 291, 361
style: *see* variation
suffixes: *n*-zero alternation 327–9; *r*-zero alternation 336–8
surnames 8–9
syllabic consistency 91
SYLLABIC UNIFORMITY 87, 211, 354
syllabification constraints: on clitics and clitic clusters 348–70
Syllable Contact Law 18, 83; in Majorcan 208–9, 210–11, 214
syllables 78–123 *see also* resyllabification

taps: alveolar 24, 30, 50; rhotics 24–5, 26, 27, 28–9, 30, 31, 32
*THREE-MANNERCODA constraint 229, 233
trills: alveolar 24, 30, 50; rhotics 24, 25, 26–7, 28, 29, 30, 31, 32–4, 39–40, 190

UNIFORMITY constraint 129n, 136, 143, 203
UNSTRESSED-LIGHT PRINCIPLE 266–7
unstressed vocoids 96–9, 115
unstressed vowels: vowel sandhi 125–32, 135–40
unstressed words 277–9

Valencia 1
Valencian 2, 3, 25, 43, 44; *alacantí* (southern Valencian) 149, 163–4; *apitxat* (*apitjat*) (pronunciation) 36; clitic clusters in 346–7; cluster reduction in 221, 226, 227, 228, 230–2, 235–49 epenthesis in 288; fricatives in 11, 13–14, 15, 20, 21, 22; stressed vowels in 38; voiced and voiceless obstruents in 149; vowel reduction in 53
variation: formal and informal 7, 179–80, 223, 228, 241, 245; *see also*; *apitxat* (*apitjat*) (pronunciation); dialects; dorso-palatal fricatives; modelling variation
velar-C clusters 177
velars: with half-open vowels 48
Veny & Pons: *Atles Lingüístic del Domini Català* 318–24
verb stress 284–8
VOICEGRADIENT constraint 160
voiced fricatives 11
voiced obstruents: distribution of 145–9; geminate 12–13
voiced plosives *see* plosives
voiced stop gemination: and Minimum Sonority Distance 265–7
voiced stops: lenition 317

voiceless obstruents *see* voiced obstruents: distribution of
vocoids 78–9, 80, 81; syllabification of 88–123 *see also* vowel sandhi
vowel epenthesis 250
vowel reduction 52–72
vowel sandhi 124–44
vowels 37–72; high vowel hiatus 91–6

western Catalan: fricatives in 14; vowel reduction in 53
word-final devoicing: non-strident obstruents 7–8
word-final *n* deletion 328
word-final sibilants 162–4
word-final voicing neutralization 149
WordMinimality constraint 120
word stress 276–84